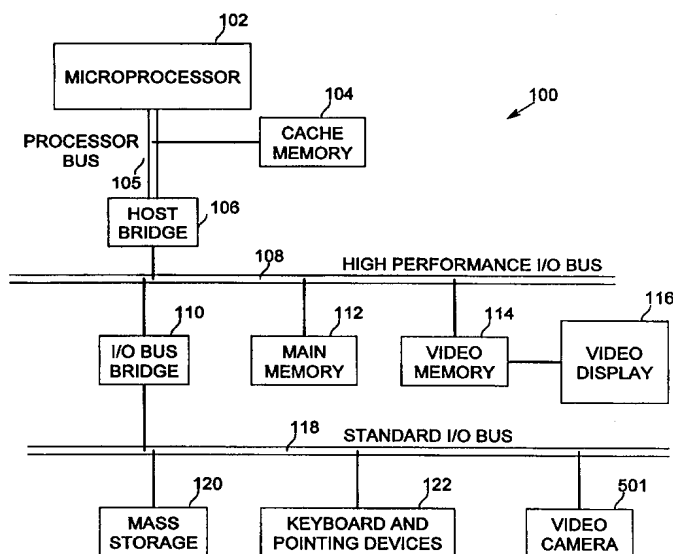




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<div>(51) International Patent Classification <sup>6</sup> : G09G 5/00</div>	<div>A1</div>	<div>(11) International Publication Number: WO 99/65014</div> <div>(43) International Publication Date: 16 December 1999 (16.12.99)</div>
<div>(21) International Application Number: PCT/US99/11542</div> <div>(22) International Filing Date: 25 May 1999 (25.05.99)</div> <div>(30) Priority Data: 09/096,282 11 June 1998 (11.06.98) US</div> <div>(71) Applicant (for all designated States except US): INTEL CORPORATION [US/US]; 2200 Mission College Boulevard, Santa Clara, CA 95052 (US).</div> <div>(72) Inventors; and</div> <div>(75) Inventors/Applicants (for US only): DELEEUEW, William, C. [US/US]; 239 N.E. 30th Avenue, Hillsboro, OR 97124 (US). KNOWLSON, Kenneth, L. [US/US]; 17289 N.W. Countryridge Drive, Portland, OR 97229 (US). NEEDHAM, Bradford, H. [US/US]; 26955 N.W. Dorland Road, Hillsboro, OR 97124 (US).</div> <div>(74) Agents: MILLIKEN, Darren, J. et al.; Blakely, Sokoloff, Taylor &amp; Zafman LLP, 7th floor, 12400 Wilshire Boulevard, Los Angeles, CA 90025 (US).</div>		<div>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</div> <div>Published With international search report.</div>

(54) Title: METHOD OF USING VIDEO REFLECTION IN PROVIDING INPUT DATA TO A COMPUTER SYSTEM



## (57) Abstract

Providing input signals to a computer system having a display (116), the computer system being coupled to a video camera (501) or other video source, is accomplished by capturing video data signals generated by the video camera (501), the video data signals representing a scene, rendering the scene on the display such that the scene is reflected and transparently visible on the display (116), analyzing the video data signals to detect an object in the scene, and generating an input signal for the computer system in response to the detected object.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<b>AL</b>	Albania	<b>ES</b>	Spain	<b>LS</b>	Lesotho	<b>SI</b>	Slovenia
<b>AM</b>	Armenia	<b>FI</b>	Finland	<b>LT</b>	Lithuania	<b>SK</b>	Slovakia
<b>AT</b>	Austria	<b>FR</b>	France	<b>LU</b>	Luxembourg	<b>SN</b>	Senegal
<b>AU</b>	Australia	<b>GA</b>	Gabon	<b>LV</b>	Latvia	<b>SZ</b>	Swaziland
<b>AZ</b>	Azerbaijan	<b>GB</b>	United Kingdom	<b>MC</b>	Monaco	<b>TD</b>	Chad
<b>BA</b>	Bosnia and Herzegovina	<b>GE</b>	Georgia	<b>MD</b>	Republic of Moldova	<b>TG</b>	Togo
<b>BB</b>	Barbados	<b>GH</b>	Ghana	<b>MG</b>	Madagascar	<b>TJ</b>	Tajikistan
<b>BE</b>	Belgium	<b>GN</b>	Guinea	<b>MK</b>	The former Yugoslav Republic of Macedonia	<b>TM</b>	Turkmenistan
<b>BF</b>	Burkina Faso	<b>GR</b>	Greece			<b>TR</b>	Turkey
<b>BG</b>	Bulgaria	<b>HU</b>	Hungary	<b>ML</b>	Mali	<b>TT</b>	Trinidad and Tobago
<b>BJ</b>	Benin	<b>IE</b>	Ireland	<b>MN</b>	Mongolia	<b>UA</b>	Ukraine
<b>BR</b>	Brazil	<b>IL</b>	Israel	<b>MR</b>	Mauritania	<b>UG</b>	Uganda
<b>BY</b>	Belarus	<b>IS</b>	Iceland	<b>MW</b>	Malawi	<b>US</b>	United States of America
<b>CA</b>	Canada	<b>IT</b>	Italy	<b>MX</b>	Mexico	<b>UZ</b>	Uzbekistan
<b>CF</b>	Central African Republic	<b>JP</b>	Japan	<b>NE</b>	Niger	<b>VN</b>	Viet Nam
<b>CG</b>	Congo	<b>KE</b>	Kenya	<b>NL</b>	Netherlands	<b>YU</b>	Yugoslavia
<b>CH</b>	Switzerland	<b>KG</b>	Kyrgyzstan	<b>NO</b>	Norway	<b>ZW</b>	Zimbabwe
<b>CI</b>	Côte d'Ivoire	<b>KP</b>	Democratic People's Republic of Korea	<b>NZ</b>	New Zealand		
<b>CM</b>	Cameroon			<b>PL</b>	Poland		
<b>CN</b>	China	<b>KR</b>	Republic of Korea	<b>PT</b>	Portugal		
<b>CU</b>	Cuba	<b>KZ</b>	Kazakstan	<b>RO</b>	Romania		
<b>CZ</b>	Czech Republic	<b>LC</b>	Saint Lucia	<b>RU</b>	Russian Federation		
<b>DE</b>	Germany	<b>LI</b>	Liechtenstein	<b>SD</b>	Sudan		
<b>DK</b>	Denmark	<b>LK</b>	Sri Lanka	<b>SE</b>	Sweden		
<b>EE</b>	Estonia	<b>LR</b>	Liberia	<b>SG</b>	Singapore		

## **METHOD OF USING VIDEO REFLECTION IN PROVIDING INPUT DATA TO A COMPUTER SYSTEM**

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

### **BACKGROUND**

#### **1. FIELD**

The present invention relates generally to computer systems and more specifically to a method for providing input data to a computer system.

#### **2. DESCRIPTION**

Many techniques and devices for communicating input data from a user to a computer system have been developed. Keyboards, cursor movement devices (such as a mouse, trackball, and joystick, for example), pens and tablets, bar code scanners, and other devices have all been used to direct a computer to perform selected tasks. With the growing use of video cameras coupled to computer systems such as personal computers (PCs), for example, opportunities for using real-time video data in new and interesting ways have arisen. However, many recent applications for video data in PCs have been focused on various display features, without regard for the possibilities of new input paradigms. What may be valuable are input techniques that take advantage of the capabilities of video cameras to enrich the PC user's experience and make PCs easier to use.

## SUMMARY

An embodiment of the present invention is a method of providing input signals to a system having a display, the system being coupled to a source of video data signals. The method includes capturing video data signals generated by the video source, the video data signals representing a scene, rendering the scene on the display such that the scene is transparently visible on the display, analyzing the video data signals to detect an object in the scene, and generating an input signal for the system in response to the detected object.

## BRIEF DESCRIPTION OF THE DRAWINGS

The features and advantages of the present invention will become apparent from the following detailed description of the present invention in which:

Figure 1 is a diagram of a sample display portion illustrating video reflection according to an embodiment of the present invention;

Figure 2 is a diagram illustrating an example of transparent graphics data displayed with operating system output graphics data according to an embodiment of the present invention;

Figure 3 is a diagram illustrating a sample computer system suitable to be programmed with a transparency method according to an embodiment of the present invention;

Figure 4 is a diagram of a software and hardware stack for implementing transparent graphics according to an embodiment of the present invention;

Figure 5 is a diagram illustrating multiple frame buffers used for providing transparent graphics according to an embodiment of the present invention;

Figure 6 is a diagram illustrating an alternating pixel technique mixing between the transparent graphics frame buffer and the operating system output frame buffer according to an embodiment of the present invention;

Figure 7 is a flow diagram for initializing a system to provide transparent graphics according to an embodiment of the present invention;

Figure 8 is a flow diagram showing double buffering control processing according to an embodiment of the present invention; and

Figure 9 is a flow diagram of color mixing and interleaving processing according to an embodiment of the present invention.

Figure 10 is a diagram illustrating components of an application program providing a video reflection input technique according to an embodiment of the present invention;

Figure 11 is a diagram of a filter graph according to an embodiment of the present invention;

Figure 12 is a flow diagram of filter initialization processing according to an embodiment of the present invention;

Figure 13 is a flow diagram of handle input events processing according to an embodiment of the present invention;

Figure 14 is a flow diagram of processing for detecting valid input events according to an embodiment of the present invention;

Figure 15 is a flow diagram describing a processing thread of a video renderer with transparency according to an embodiment of the present invention; and

Figure 16 is a flow diagram describing a processing thread of a transparency mixer core according to an embodiment of the present invention.

## DETAILED DESCRIPTION

In the following description, various aspects of the present invention will be described. For purposes of explanation, specific numbers, systems and configurations are set

forth in order to provide a thorough understanding of the present invention. However, it will also be apparent to one skilled in the art that the present invention may be practiced without the specific details. In other instances, well known features are omitted or simplified in order not to obscure the present invention.

An embodiment of the present invention is a method for providing user input data signals to a computer system such as a PC, for example. Real-time video images may be captured by a video camera coupled to a PC or received from any other source of video data signals and rendered to the entire screen of the PC's display such that the captured video images may be made transparent and shown on the display. The video data signals may be mixed with the normal application program and operating system software display output data signals so as not to obscure the normal foreground display images. The resulting effect is that when the video camera is pointed toward the user, the user may see a reflected image of himself or herself on the computer screen while the user may also be able to see and interact with other display elements such as desktop icons or application program displays. Because the reflected image is displayed transparently, it does not interfere with the display of the other display elements.

Although some embodiments of the present invention are described herein with reference to a video camera, the invention is not so limited in this respect. In alternate embodiments, the source of the video data signals may be any video source, such as a video cassette recorder (VCR), broadcast television signals, or previously stored video data signals, for example.

One of the features of some embodiments of the present invention is that the user may see his or her reflection or image on the computer screen after it is captured by the video camera. Using this video reflection, the user may interact with application program and operating system display elements of the computer system's graphical user interface (GUI) by physically moving real objects (such as the user's hands and fingers, for example) that are in the field of view of the video camera. The input capability of some embodiments of the present invention may be provided without any modification to existing application

programs. In embodiments of the present invention, these movements may be used to direct actions within the computer system such as grasping or selecting icons and other application program elements, much as the user now does with a mouse or other selection device. This physical movement activity to create a desired result while observing the user's reflected image is very natural and intuitive for users. For example, users frequently look at reflected images of themselves in a mirror, for example, while manipulating objects with their hands without thinking very much about it. Additionally, it is typically easy for users to use either a dominant or non-dominant hand or both hands to perform operations while viewing their reflected image. In alternate embodiments, specific predetermined props may be used as the objects to direct actions within the computer. These props may be active or inactive. Active props contain mechanical or electrical mechanisms such as light emitting diode (LED) lights. Inactive props are items such as colored sticks or paddles, fabric swatches, and colored gloves, for example, although the present invention is clearly not limited in this respect.

While some embodiments of the present invention may not replace the use of a computer mouse entirely, it is a powerful new paradigm for user input in a computer system. In particular, it presents a natural and direct method for manipulating computer display objects by moving real objects. Various types of application programs may be extended to make beneficial use of this input paradigm. Computer games, for example, may employ user input through video reflection for increased user interaction. Video presentation systems wherein a user points to selected areas of video display data may be made easier to use. Video-conferencing applications may be extended to provide additional input features. A combination PC/television (TV) may make use of this input capability, especially where the distance between the user and display is relatively large. For many users, such as children, for example, the manipulation of real objects (such as their hands) to send input to a computer system is more intuitive than other typical input devices.

In one embodiment, the colors of pixels in the video data signals sent to the PC for processing and display may be analyzed to track the movement of real objects such as a user's hands. This may be assisted through the use of colored dots attached to the user's

hands or fingers or through other computer vision methods. Optionally, distinctive colors may be placed on specific objects (like fingers and thumbs, for example, or other props), to help the computer vision components of embodiments of the present invention isolate and identify objects to be recognized as input event generators.

In one embodiment, the cursor of the PC's mouse may be moved in accordance with the movement of a user's index finger and thumb. When the user "pinches" his or her index finger and thumb and the images of the index finger and thumb are detected in the captured video images, a mouse down event may be introduced to the operating system software. When the user releases the pinched hand position and these images of the index finger and thumb are detected in the captured video images, a mouse up event may be introduced. Other movements signifying specific input signals and events may also be defined in alternate embodiments.

Figure 1 is a diagram of a sample display portion illustrating video reflection according to an embodiment of the present invention. In this example, the image of the user is captured by the video camera and rendered in a transparent manner to the display. The user then interacts with display objects, such as application program icons, for example, to provide input signals and events to the system.

An underlying capability of some embodiments of the present invention is a method of providing a transparent layer of display data signals (such as video data signals communicated by a video camera, for example) over the top of another layer of display data signals on a computer display so that the user may see both layers clearly and substantially simultaneously. This capability will be described first for general use and further below for use with an input detection method of some embodiments of the present invention. The capability to display transparent windows doubles, in essence, the maximum screen area available on a display for use by application programs. One embodiment is a method for producing transparent computer graphics layers by interleaving (or alternating in a pattern) the pixels from one video frame buffer with the pixels from another video frame buffer. In this embodiment, selected pixels from a first frame buffer are mixed by color averaging with

corresponding pixels from a second frame buffer to reduce the “checkerboard” effect created by the use of spatial multiplexing alone. Additionally, because the degree of interleaving is adjustable and the color averaging may be weighted, the degree of transparency of the displayed images may be controlled.

In this embodiment, an output frame buffer used by operating system software is not affected by provision of the transparency feature and the operating system is unaware of the transparency operations. Hence, the transparency effect provided by embodiments of the present invention does not require modifications to application programs for transparency to work over them. Furthermore, input operations to application program and operating system windows are not affected by transparent foreground effects.

An embodiment of the present invention operates by combining at least two frame buffers of computer graphics output data or video data in the form of electrical signals. The pixels of the output, or visible, frame buffer are created by spatially interleaving the contents of two input frame buffers. The interleaving in this embodiment is accomplished by combining pixels of one frame buffer with those of the other frame buffer. This results in the visual illusion of two displays of images layered one on another. As the pixels are being interleaved, the pixels of the first frame buffer are color averaged with the pixels of the second frame buffer that they are about to replace. Color averaging is performed on the pixels of one frame buffer by averaging them with the corresponding pixels of the other frame buffer prior to, or during, interleaving them into the output frame buffer. The result comprises multiple overlapping images being substantially simultaneously visible on a display such as a computer monitor, for example.

Figure 2 is a diagram illustrating an example of transparent graphics data displayed with operating system output graphics data according to an embodiment of the present invention. Operating system output frame buffer 10 is an area of memory used to store the current display data of the computer system shown below in Figure 3. The operating system output frame buffer may be allocated in any memory available to the operating system. A frame buffer is a set of storage locations to store a two-dimensional array of pixel data. The

operating system output frame buffer may be associated with operating system software of the computer system, which controls the generation and display of the data signals on a computer monitor (not shown). In one embodiment, the operating system software comprises the Windows 95® or Windows NT® operating system software available from Microsoft Corporation, although other operating system software supporting graphical user interfaces may also be employed. In this example, the operating system output frame buffer 10 contains application program display data signals for three overlapping windows shown pictorially in Figure 2 and labeled 12, 14, and 16, respectively.

Transparent graphics frame buffer 18 is an area of memory used to store the display data of transparent graphics for substantially simultaneous display with the display data signals of the operating system output frame buffer. This area of memory may be allocated in any memory available in the system. Transparent graphics frame buffer 18 may be used to store frames of video data received from a video camera or other source of video data signals. In the example shown in Figure 2, display components such as a clock 20 and stock ticker 22 are shown as sample application program display features which illustrate the use of transparency, although generally any display components may be made transparent through the use of embodiments of the present invention. More specifically, in embodiments of the present invention, captured video frames may be shown as a sequence of transparent images on the display.

The display components of the operating system output frame buffer and the transparent graphics frame buffer may be combined in block 24 by color mixing selected corresponding pixels of each buffer while interleaving the resulting pixels of the color mixing operation with the operating system output frame buffer's pixels to form the display components of visible display buffer 28. The visible display buffer shows in pictorial form the three overlapping windows 12, 14, and 16 with the clock 20 and stock ticker 22 displays appearing as transparent display components overlaying portions of the windows. In this example, the transparent display components are partially overlaying the other display components. However, it should be understood that the transparent display components may

be entirely within the boundaries of one or more non-transparent windows or display components on the display. Of course, in certain application programs and with certain display components, the display of data from two display components with one substantially or even completely on top of the other may present image quality problems for the user. Nonetheless, in other application programs the ability to overlay transparent display components in a well designed manner is advantageous and desirable.

In addition, some embodiments of the present invention allow transparent display components overlaying background windows to have little or no effect on input operations to a selected background window. For example, a user may interact with an input window of an application program being displayed in a background window while a transparent display component is partially or completely overlaying the background window. Generally, the operating system software may accept the user input events or key strikes to the input window (such as a mouse entry or text entry) without substantial interference with the display of the transparent display components.

In accordance with some embodiments of the present invention, a method for producing transparency effects employs minimal mixing of display contents. Instead, it relies on the human eye's inability to distinguish between the color of adjacent pixels on a computer monitor (in essence, the human eye averages each pixel with its neighbor). Some mixing is employed, because large computer monitors and low display resolutions may result in a "checkerboard" effect when pixels are interleaved in this manner. In one embodiment, one half of the pixels from a first frame buffer (such as the operating system output frame buffer) are averaged with the corresponding one half of the pixels from a second frame buffer (such as the transparent graphics frame buffer) as the pixels of the two frame buffers are interleaved into a third display buffer. By averaging a fraction of the pixels, there may be a decrease in the processing power used when providing the transparency effect. In alternate embodiments, different percentages of pixels may be averaged (e.g., one fourth of the pixels, one eighth of the pixels, one sixteenth of the pixels, one thirty-second of the pixels, or any

one Nth of the pixels where N is a positive integer), and the percentages may be changed dynamically

Figure 3 is a diagram illustrating a sample computer system suitable to be programmed according to an embodiment of a method for producing transparency displays in conjunction with video reflection to obtain user input signals in accordance with the present invention. Sample system 100 may be used, for example, to execute the processing for the methods described herein. Sample system 100 is representative of computer systems based on the PENTIUM®, PENTIUM® Pro, and PENTIUM® II microprocessors available from Intel Corporation, although other systems (including personal computers (PCs) having other microprocessors, engineering workstations, set-top boxes and the like) may also be used. Sample system 100 includes microprocessor 102 and cache memory 104 coupled to each other through processor bus 105. Sample system 100 also includes high performance I/O bus 108 and standard I/O bus 118. Processor bus 105 and high performance I/O bus 108 are bridged by host bridge 106, whereas high performance I/O bus 108 and standard I/O bus 118 are bridged by I/O bus bridge 110. Coupled to high performance I/O bus 108 are main memory 112 and video memory 114. Coupled to video memory 114 is video display 116. Coupled to standard I/O bus 118 are mass storage 120, and keyboard and pointing devices 122. In one embodiment, a video camera 501 may also be coupled to standard I/O bus 118.

These elements perform their conventional functions well-known in the art. In particular, mass storage 120 may be used to provide long-term storage for the executable instructions for embodiments of methods for obtaining user input signals through the use of video reflection in accordance with the present invention, whereas main memory 112 is used to store on a shorter term basis the executable instructions of embodiments of the methods for obtaining user input signals through the use of video reflection in accordance with the present invention during execution by microprocessor 102.

Figure 4 is a diagram of a software and hardware stack for implementing transparent graphics according to an embodiment of the present invention. Application programs 200 designed to use transparent display objects call functions provided by transparency support

software 202 to define and update the transparent display objects. In response, transparency support 202 calls the operating system graphics rendering programming interface (graphics API) 204 in this embodiment. In the Windows95® operating system, this may be the Graphics Device Interface (GDI). The transparency support software 202 also calls the operating system's video hardware control abstraction programming interface (video control API) 206 in this embodiment. In the Windows95® operating system, this may be the DirectDraw™ API, available from Microsoft Corporation. In some operating systems, the graphics API 204 and video control API 206 may not be distinguishable from each other as they may exist within the same application programming interface. The graphics API 204 may be used to render requested graphics to the transparent graphics frame buffer 18 shown in Figure 2. The video control API 206 may be used to control frame buffer visibility and to access the contents of all frame buffers. In this embodiment, the graphics API 204 and video control API 206 interact with display driver software 208 to communicate with video card 210. The video card 210 controls the video display 116 in the system of Figure 3. Video card accesses video memory 114 to obtain display data. Other application programs 212, which do not employ transparency, interact with the graphics API 204 to create and update display objects.

Generally, images may be displayed on a display such as a computer monitor, for example, by creating a frame buffer of pixel data in video memory 114. This frame buffer may be designated as a visible portion of video memory by video control API 206. If there is a sufficient amount of video memory available, multiple frame buffers may be defined, only one of which may be used at a time (by the video card 210) to obtain the data signals for building the current visible display. In a well-known double buffering technique, a first frame buffer is considered to be the "visible" buffer and the video card 210 reads data signals from it to obtain the current display data signals, while a second frame buffer (or "non-visible" buffer) is written to with new display data. In this embodiment, the video control API is then called upon to "flip" the frame buffers by designating the second frame buffer to be the visible buffer and designating the first frame buffer to be the non-visible buffer. Use

of this technique provides for the smooth update of display data, resulting in aesthetically pleasing displays for the user. Embodiments of the present invention may extend this concept to employ extra frame buffers to provide the transparent display data signals in conjunction with normal display data.

Figure 5 is a diagram illustrating an embodiment of multiple frame buffers used for providing transparent graphics. At least one designated portion of the video memory may be assigned to be displayed as visible on the computer monitor at a time. This is called the "visible display". That is, the visible display comprises the display data from an area of video memory that is currently displayed on the computer monitor for viewing by a user. Generally, in this embodiment the graphics API 204 of the operating system software writes data signals into the operating system output frame buffer 10. In most current systems, the operating system output frame buffer, resident in video memory 114, is used for the visible display. However, in embodiments of the present invention, other frame buffers may be used as the visible display. A first working frame buffer 300 and a second working frame buffer 302, both resident in video memory 114 or other accessible memory, store display data according to embodiments of the present invention. In this embodiment, each frame buffer stores an array of pixel data signals. The size of the array in this embodiment is dependent on the current display characteristics of the system. Frame buffer array sizes may, for example, be 640 pixels by 480 pixels, 800 pixels by 600 pixels, or 1280 pixels by 1024 pixels, or other appropriate sizes dependent on the computer monitor and operating system software settings. Each pixel includes red (R), green (G), blue (B), and optionally, opacity (A) components. Alternatively, other color coding schemes such as YUV or YUVA may also be used. Multiple transparent graphics frame buffers 18, 19, resident in main memory 112 or in video memory 114, in this embodiment store transparent display data signals created by transparency support software 202, video control API 206, and graphics API 204. The transparent graphics frame buffers 18, 19 are "double-buffered" similar to the working frame buffers to provide rapid updates to the transparently displayed images.

In one embodiment, data signals from one of the transparent graphics frame buffers 18, 19 may be color mixed and interleaved with data signals from operating system output frame buffer 10, and then stored in one of the working frame buffers. In an alternate embodiment, the mixing may be accomplished with the assistance of dedicated circuitry for alpha blending. This mixed and interleaved data may be stored into a working frame buffer when the working frame buffer is in a "non-visible" state (that is, in this embodiment the data stored in the frame buffer is not currently displayed on the computer monitor). While one of the working frame buffers is being written to in a non-visible state, the other working frame buffer may be in a "visible" state and used as the source of current display data. When the color mixing and interleaving operations are complete for a working frame buffer, the non-visible working frame buffer may be designated the visible working frame buffer and vice versa. This double buffering of the working frame buffers process may be repeated at a rate of at least 8-15 times per second in this embodiment to provide a visually appealing display to a user. Additionally, in this embodiment use of the transparent graphics frame buffers is also double buffered. In this embodiment, double buffering of the transparent graphics frame buffers is employed because color mixing may not be performed on a frame buffer that is currently being modified, and a frame buffer that is being mixed from may not be updated at the same time. Without double buffering, errors may result in the displayed images or unnecessary processing delays may be incurred.

In some embodiments of the present invention, interleaving of the pixels of one of the transparent graphics frame buffers and the operating system output frame buffer may be accomplished as follows. In one embodiment, alternating pixels in the selected working frame buffer may be written by a mix of a transparent graphics frame buffer pixel value and a spatially corresponding operating system output frame buffer pixel value. The other pixels in the selected working frame buffer may be written with pixels from the operating system output frame buffer. In another embodiment, pixels from the operating system output frame buffer may be block transferred to the selected working frame buffer and pixels from the transparent graphics frame buffer may be subsequently spatially multiplexed and color

averaged with the pixels of the selected working frame buffer. In another embodiment, the mixing may be accomplished with the assistance of circuitry on the video card for performing alpha blending. Note that interleaving may not be necessary when using dedicated circuitry for alpha blending.

Figure 6 is a diagram illustrating an embodiment of one method of alternating pixel mixing between one of the transparent graphics frame buffers and the operating system output frame buffer. A "T+OS Mix" pixel in the selected working frame buffer comprises a color averaged mix of a pixel from the selected transparent graphics frame buffer (the T value) and a pixel from the operating system output frame buffer (the OS value). An "OS" pixel in the selected working frame buffer contains a spatially corresponding pixel copied from the operating system output frame buffer. In this embodiment, color averaging may be performed through a weighted averaging scheme on each color component of each pixel from corresponding positions within the two frame buffers, although in other embodiments, different color mixing techniques may also be employed. In one embodiment, weighted averaging may be accomplished by multiplying a component value of a first pixel by a weight value and multiplying the same component value of a second pixel by a different weight value. The two weighted color components may then be added together and the resulting sum may be divided by the sum of the two weight values. This method is also known as alpha blending. By using this alternating pattern, the processing employed to create the transparent effect may be cut in half, for example, in comparison to a mixing of all pixels of the frame buffers. The pixel data movement within the video memory may be performed by a block transfer operation provided by the drawing API or video control API in this embodiment.

In other embodiments, the mixed pixels may comprise only one quarter of the pixels in the selected working frame buffer, one eighth of the pixels in the selected working frame buffer, or other percentages such as any one Nth of the pixels, where N is a positive integer, depending on the specific interleaving pattern used. Furthermore, in other embodiments the interleaving pattern may be modified. For example, the interleaving pattern may comprise

horizontally alternating lines from the transparent graphics frame buffer and the operating system software frame buffer. Alternatively, the interleaving pattern may comprise vertically alternating lines from the transparent graphics frame buffer and the operating system software frame buffer. A combination of a checkerboard pattern and horizontally or vertically alternating lines may also be used. One skilled in the art will realize that various interleaving patterns may be used in embodiments of the present invention with varying degrees of transparent effect, and the invention is not limited in scope to any particular pattern.

In another embodiment of the present invention, the interleaving pattern may be changed over time at a periodic or non-periodic rate or in a predetermined manner. For example, use of any two of the different interleaving patterns described above may be alternated, such that a first interleaving pattern is used for a first generation of the transparent graphics frame buffer and a second interleaving pattern is used for a second, succeeding generation of the transparent graphics frame buffer. This process may be repeated, thereby implementing a hybrid spatial, color-mixed, and temporal transparency method.

In another embodiment, special purpose circuitry may be used to eliminate the need to mix the pixels in software.

It should be noted that each pixel in the transparent frame buffer may be used more than once, or not at all, to achieve a stretching or shrinking effect in the resulting transparency output. The frequency and location of pixel re-use or omission depends at least in part on the desired amount of stretching or shrinking.

Figure 7 is a flow diagram illustrating an embodiment for initializing a system to provide transparent graphics. At block 400, the operating system display output control information may be determined. This control information comprises the size of the display, color resolution, and other data. Next, at block 402, two working frame buffers may be allocated in video memory in this embodiment. These operations may be performed by calls to the video control API in this embodiment. At block 404, a block transfer operation may be performed to copy data from the normally visible operating system output frame buffer to a selected one of the two working frame buffers. Assume for this example that the second

working frame buffer is selected first, although the first working frame buffer may also be used as the initial working frame buffer. The block transfer may be performed by a call to the video control API in this embodiment. At block 406, the operating system output frame buffer may be set to a "non-visible" state by a call to the video control API. At block 408, the selected working frame buffer (for example, the second working frame buffer) may be made visible by a call to the video control API in this embodiment. In some embodiments, block 406 and block 408 may be accomplished by a single call to the video control API. At this point, the video card's current display output data may be obtained from the selected working frame buffer, not the operating system output frame buffer. In alternate embodiments, other APIs may also be used to effect the same results.

Figure 8 is a flow diagram showing an embodiment of double buffering control processing. After starting block 410, a block transfer operation may be performed at block 412 to copy the operating system output frame buffer to the non-visible first working frame buffer by a call to the video control API in this embodiment. At block 414, an operation may be performed to write the mixed and interleaved contents of the first working frame buffer and a selected one of the transparent graphics frame buffers to the first working frame buffer. At block 416, the first working frame buffer may be made visible and the second working frame buffer may be made non-visible, in effect, flipping the two frame buffers as the current display output data source. At block 418, a block transfer operation may be performed to copy the operating system output frame buffer to the non-visible second working frame buffer by a call to the video control API in this embodiment. At block 420, an operation may be performed to write the color mixed and interleaved contents of the second working frame buffer and a selected one of the transparent graphics frame buffers to the second working frame buffer. The selected transparent graphics frame buffer may be the transparent graphics frame buffer previously selected at block 414 or the other transparent graphics frame buffer. At block 422, the second working frame buffer may be made visible and the first working frame buffer may be made non-visible, in effect, flipping the two frame buffers as the current display output data source. This process may be repeated by returning to block 412. During

each of the previous blocks, the operating system software may be concurrently writing additional display data into the operating system output frame buffer.

an embodiment of the color mixing and interleaving operation of blocks 414 and 420 is further described with reference to Figure 9. At block 426, a memory location in the currently non-visible (either the first or the second) working frame buffer may be determined for a reference point (e.g., point M 304) of the currently selected transparent graphics frame buffer. At block 428, a data signal value for a pixel from the currently non-visible working frame buffer may be read and the spatially corresponding pixel(s) from the currently selected transparent graphics frame buffer may be determined. This correspondence may not necessarily be a 1:1 ratio since the transparent graphics frame buffer image may be stretched or reduced to fit a portion of the working frame buffer. This pixel correspondence determination is well-known in the art and is commonly used in stretch block transfers in operating system software (e.g., the StretchBlt function in the Windows95® operation system). Next, at block 430, in this embodiment the weighted average of the pixel from the working frame buffer and the pixel from the currently selected transparent graphics frame buffer may be computed. The weighted averages of the individual pixel components may be determined on a color component by color component basis. That is, red components may be averaged, blue components may be averaged, and green components may be averaged. The weight that is given to each of the components determines the resulting transparency of the pixel, however the same weight value may be used for all components of a given pixel. It is the weight associated with a pixel that affects at least in part the level of transparency. These weights may be manipulated by the application program employing transparency to achieve various mixing ratios. Furthermore, the application program employing transparency may provide user interface elements that allow the user to control the mixing ratios directly or indirectly.

The result of the weighted averaging computation may be placed into the same location in the working frame buffer at block 432 as the current pixel being processed. At block 434, the next location in the working frame buffer to be processed may be determined,

taking into account the current interleaving pattern (e. g., using every second pixel, every fourth pixel, horizontally or vertically alternating lines, etc.). At block 436, if more pixels of the working frame buffer and the currently selected transparent graphics frame buffer are to be processed, processing continues with block 428 with the next pixel. Otherwise, color mixing and interleaving processing ends at block 438.

In an alternate embodiment, mixing of pixels may be performed through a video control API block transfer using dedicated alpha blending circuitry.

Figure 10 is a diagram illustrating components of an application program providing a video reflection input technique according to an embodiment of the present invention. Video data signals 500 may be generated by a video camera 501 coupled to a computer system such as is shown in Figure 3. Or by any other video source. The video camera may be coupled to the standard I/O bus 118 and provides the video data signals 500 for storage in mass storage 120 or main memory 112. Once the video data signals are within the computer system 100, they may be processed by programs being executed by microprocessor 102, such as application program 502 and filter graph manager 504 of Figure 10, for example. Referring back to Figure 10, application program 502 may be designed to provide any desired functionality for a user. At a minimum, in this embodiment the application program is designed to provide the capability of user input through the use of video reflection. The application program may interact with other application programs (not shown) and the operating system software (not shown) being executed by the processor of the computer system. When application program 502 is designed to provide the capability of user input selections through the use of video reflection in accordance with some embodiments of the present invention, the application program comprises filter graph manager 504, filter initialization 506, handle input events 508, filter graph 510, and COM interfaces 512.

In one embodiment, a filter graph 510 may be created as part of a filter graph architecture of the ActiveMovie™ application programming interface (API) (also known as the DirectShow™ or DirectMedia™ API) available from Microsoft Corporation, although in other embodiments, other APIs may be used and the invention is not limited in scope in this

respect. ActiveMovie™ is a software architecture that controls and processes streams of multimedia data signals, as well as run-time environment software that uses the architecture to allow users to play digital video and sound encoded in various well-known data formats on a PC. The playback capability makes use of video and audio hardware cards that support the DirectX™ set of APIs, available from Microsoft Corporation, although other APIs may also be used and the invention is not limited in scope in this respect. The ActiveMovie™ architecture defines how to control and process streams of time-stamped multimedia data signals by using modular software components called filters connected in a configuration called a filter graph 510.

A software object called a filter graph manager 504 may be accessed by the application program 502 to control how the filter graph 510 may be assembled and how data may be moved through the filter graph. The instantiated filter graph manager generates and manages the filter graph. The filter graph comprises at least one custom filter for processing video data signals. The filter graph manager 504 provides a set of Component Object Model (COM) interfaces 512 to allow communication between the filter graph and application program 502. Application program 502 may call the filter graph manager COM interfaces to control processing of the multimedia data stream or retrieve filter events. Since application program 502 processes the incoming video data stream from the coupled video camera 501, in this embodiment the filter graph manager and filter graph are incorporated into the application program. In other embodiments, the filter graph manager may be a separate program independent of the application program.

In one embodiment, all components of the filter graph architecture may be implemented as COM objects. This includes the filters through which data signals may be passed, and filter components that serve as connections between filters or that allocate memory. Each object implements one or more interfaces, each of which comprises a predefined set of functions called methods. Generally, application program 502 calls a method to communicate with the object exposing the interface. For example, the application program may call methods on an IMediaControl interface on the object of filter graph

manager 504, such as a Run method, which starts a multimedia data stream. Filter graph manager, in turn, calls a Run method on the IFilter interface exposed by each of the filters of the filter graph 510.

A filter graph 510 may comprise a collection of filters of different types. Most filters may be categorized as one of three types. A source filter takes data signals from some source, such as a file on disk, a satellite feed, an Internet server, or a video camera 501, and introduces the data signals into the filter graph. A transform filter takes the data signals, processes them, and passes them along to another filter. A rendering filter renders the data signals. Typically, the processed data signals may be rendered to a hardware device such as a computer display, but the data may also be rendered to any location that accepts media input (such as a memory or a disk file, for example).

Generally, filter graphs may be used to stream multimedia data, such as video data, for example, through filters. In a multimedia data stream, one filter passes the multimedia data downstream to the next filter in the filter graph. To make a filter graph operate, filters may be connected in a predetermined order, and the multimedia data stream may be started and stopped in the predetermined order. The filter graph manager connects filters and controls the multimedia data stream. Controlling the multimedia data stream comprises starting, pausing, or stopping the multimedia data stream, playing the data stream for a particular duration, or seeking a particular point in the data stream. The filter graph manager 504 allows the application program 502 to specify these activities, and then calls the appropriate methods on the filters to invoke them. The filter graph manager also allows filters to post event notifications that the application program may retrieve. In this way, the application program may retrieve the status of a filter that the application program has installed in the filter graph, once multimedia data is flowing through the filter graph.

During processing initialization, application program 502 instructs the filter graph manager 504 to load a graph file (not shown) which describes the filter graph. The filter graph manager creates the filter graph 510 based on the graph file and filter behavior commands received from filter initialization 506. The filter behavior commands refer to

specific filters by name and specify input and control parameters for the filters. When video data 500 is streaming through the filter graph and being processed according to specified filters, event notifications (also known as window messages) may be sent from the filter graph through COM interfaces 512 to handle input events 508. In one embodiment, the input event may be determined to be a mouse down event or a mouse up event. In other embodiments, the input event may be any user input selection detected during the filter graph's processing of the video data stream. Handle input events 508 receives the event notification and handles it for application program 502. In some embodiments, the handle input events 508 may comprise calls to other components within the application program to perform desired functions.

Figure 11 is a diagram of a filter graph according to an embodiment of the present invention. In this embodiment, the filter graph comprises seven filters: video capture filter 520, tee filter 522, optional first and second color space converter filters 524, 526, video renderer with transparency filter 528, color analyzer filter 530, and blob detector 532, although the invention is not limited in this respect. Video capture filter 520 captures individual frames of video data signals received by the PC from a coupled video camera or other video source. Each frame of video data signals may be passed to tee filter 522. The tee filter splits the video data signals into two substantially similar data streams. One of the data streams may be forwarded to first color space converter filter 524, and the other data stream may be forwarded to second color space converter 526. The first color space converter converts the first data stream such that the color space of the video data signals is similar to the system display color space. In one embodiment, the pixel format of the video data signals must match the video card pixel format (e.g., 16 bit RGB color or 24 bit RGB color, for example). Converted data signals from the first color space converter filter may be forwarded to the video renderer with transparency filter 528. The video renderer with transparency filter 528 renders the first data stream on a display in a transparent manner as described above with respect to Figures 4 through 9. Video data signals may also be converted or "flipped" about a vertical center line of each frame to make objects on one side

of the scene appear on the same side of the display (in effect, the scene appears like a mirror on the display).

In this embodiment, the second color space converter filter 526 converts the second data stream into a red, green, blue, alpha (RGBA) color format for each pixel of the second data stream, wherein the alpha values may be used to hold extra pixel information. Converted data signals from the second color space converter filter may be forwarded to the color analyzer filter 530. The color analyzer filter analyzes the pixels of the second data stream and tags color values of pixels within specified color ranges with specific alpha "A" values. The second data stream with added tag information may be forwarded to blob detector filter 532.

The blob detector filter analyzes the data stream and finds blocks of pixels with specific alpha values and posts event notifications (e.g., messages) to a control window of the application program 502. The application program forwards these event notifications to the handle input events function 508. In one embodiment, the blob detector filter may be set to look for substantially contiguous blocks of a predetermined color that have been tagged by the color analyzer filter. A block is also known as a blob of color. Herein, a blob may have predetermined physical attributes with a distinctive color and may be represented digitally as a group of pixels. The blob detector filter may be informed of a specified window handle to which detected blobs may be reported. In one embodiment, a blob of a predetermined pattern is determined to be an input event when the blob is detected in the video data stream by the blob detector filter and the handle input events function verifies that the blob meets a set or predetermined criteria. The location of the blobs in the video data stream (that is, in the current video frame) may also be communicated to handle input events 508.

In the embodiment shown in Figure 11, the blob detector filter and the color analyzer filter are shown as separate filters, although in alternate embodiments the functions of these filters may be combined into a single filter, or may be combined with the video renderer filter. One skilled in the art will recognize that various combinations of such filters may be used, and the present invention is not limited in this respect.

Figure 12 is a flow diagram of filter initialization processing according to an embodiment of the present invention. Generally, filter initialization 506 instructs the filter graph manager 504 to create and initialize the filter graph 510 and begin the flow of video data signals into the filter graph. At block 540, filter initialization directs the filter graph manager to create a filter graph. At block 542, the video renderer with transparency filter may be located in the filter graph. Locating a filter means that an IFilter COM interface is retrieved based on the filter name and a configuration interface is then obtained through the filter. Next, at block 544, initialization parameters for the video renderer with transparency filter may be set. For example, the desired level of opacity may be set for future video data signal processing. Other initialization parameters include mirroring settings and interleaving patterns. At block 546, the color analyzer filter may be located in the filter graph. At block 548, the color ranges and corresponding alpha values of pixels in the video data stream to be tracked by the color analyzer filter may be set. For example, in the embodiment tracking the finger and thumb input events, the color ranges and alpha values of the finger and thumb may be set so that the color analyzer tags pixels in the video data stream having color values in the specified ranges. At block 550, the blob detector filter may be located in the filter graph. At block 552, a control window identifier, alpha values, and corresponding control messages may be set in the blob detector filter. At block 554, video data signal flow through the filter graph may be started.

Once the video data signals are flowing through the filter graph, an event notification called a window message may be received by handle input events 508 from the blob detector filter 532. Figure 13 is a flow diagram of handle input events processing according to an embodiment of the present invention. Processing begins by waiting for a window message ready indication at block 560. If no window message ready event is received, then handle input events continues to wait until a potential input event is received. When a window message is received, further processing may be performed to determine if it is a potential input event. At block 562, the window message may be examined to determine if a blob type 1 event has been detected by the blob detector filter. In one embodiment, a blob type 1 may

be associated with a predetermined physical artifact occurring in a video frame such as an index finger, for example, and a blob type 2 may be associated with a another physical artifact occurring in a video frame such as a thumb, for example, although one skilled in the art will recognize that various physical artifacts and objects and their accompanying colors may be used in various embodiments as input indicators. In this embodiment, two distinct colors may be used as indicators, however in other embodiments any number of colors may be employed. In this embodiment, if blob type 1 is detected, then the size and location of the blob type 1 in the current video frame of the video data stream may be recorded in a data structure denoted the type 1 list at block 564. Processing then continues by waiting for another potential input event.

If blob type 1 was not detected at block 562, then processing continues at block 566, wherein the window message may be examined to determine if a blob type 2 event has been detected by the blob detector filter 532. If blob type 2 was detected, then the size and location of the blob type 2 in the current video frame of the video data stream may be recorded in a data structure denoted the type 2 list at block 568. Processing then continues by waiting for another potential input event. If blob type 2 was not detected, processing continues at block 570, wherein the window message may be examined to determine if an end of frame indicator has been detected by the blob detector filter. If an end of frame indicator was detected, then block 572 is processed to determine if a valid input event was detected in the current video frame. After block 572, processing continues by waiting for another window message at block 560. When the end of frame indicator is detected, the blobs detected during processing of the frame are stored in the blob lists. In one embodiment, the type 1 and type 2 lists store possible occurrences of two different physical objects represented in the video data stream. In other embodiments, any number of lists may be used to track a corresponding number of physical objects represented in the video data stream.

If an end of frame indicator was not detected, the window message may be examined at block 574 to determine if an application close message was received. If an application

close message was received, the filter graph may be stopped at block 576, the filter graph may be deleted at block 578, and the application program 502 may be exited at block 580. Otherwise, an application close message was not received and processing continues by waiting for a window message at block 560.

Figure 14 is a flow diagram of processing for detecting valid input events. At the beginning of detecting valid input events processing, a complete video frame may be processed by the filter graph and application program 502 may have at least two lists of blobs occurring in the video frame. In one embodiment, the two lists are the blob type 1 and blob type 2 lists, each list representing a blob color region space. Valid input event detection processing then comprises searching the blob lists to find the closest pair of blobs. Each pair is tested to determine if it is closer together than a selected threshold distance. If such a blob pair exists, then detection of this blob pair is considered to be a valid input event. If more than one pair is detected, then the closest pair is used for further processing. At block 600, detect valid input processing determines if the end of the blob type 1 list has been reached. If the end of the blob type 1 list has not been reached, then block 602 is performed to find a blob in the blob type 2 list that is closest in the video frame to the current blob type 1 and to record the index within the blob type 2 list of this closest blob type 2 into the current entry of the blob type 1 list. In this embodiment, closeness may be defined as the spatial distance between the centers of two blob locations in the two dimensional plane of the video frame. At block 604, the next entry in the blob type 1 list is selected for processing. Processing continues with this next blob type 1 at block 600. If there is no pair available at 605, then an appropriate flag or other indicator is set in the current entry of the blob type 1 list.

When the end of the blob type 1 list is reached, block 606 may be performed to find the closest blob type 1/blob type 2 pair in the blob type 1 list. If the closest pair is less than a threshold distance apart at block 608, then block 610 may be performed to submit the detected input event to the operating system software of the computer system. The operating system software may then pass the input event to other application programs as needed. In one embodiment, the location of the blobs in the current video frame may be mapped or

otherwise converted to screen coordinates (e.g., normalized mouse coordinates) prior to reporting the input event. At block 610, if the closest pair is not less than the threshold distance apart, then block 611 may be performed to submit a different input event to the operating system software. Next, at block 612 the type 1 blob list and the type 2 blob list may be reset. Detect valid input event processing ends at block 614.

While the example of handling input events by detecting the movement of a user's finger and thumb has been described, one skilled in the art will recognize that the capture and analysis of a video data stream including a number of different physical objects of different colors and movements may be used to signify an input event, and that the invention is not limited in scope by way of the specific example shown herein. Additionally, in this embodiment, lists are used to store the blob information. In other embodiments, other data structures may be used and other techniques for searching the data structures to determine blob pairs may be also be used.

Figure 15 is a flow diagram describing a processing thread of a video renderer with transparency according to an embodiment of the present invention. Video renderer with transparency processing begins at block 700. At block 702, a video control API object (such as a DirectDraw™ object, for example), three video frame buffers and a secondary thread may be created. At block 704, if the transparent graphics frame buffers do not exist or one or more of their sizes have changed, then new transparent graphics frame buffers may be created and a first one of the transparent graphics frame buffers may be marked as the currently active frame buffer. At block 706, a graphics display interface (GDI) display context (DC) for the currently active transparent graphics frame buffer may be created. At block 708, the GDI may be used by the video renderer to draw video data signals on the transparent DC created in block 706. Next, at block 710 the currently inactive transparent graphics frame buffer may be marked as active and the active transparent graphics frame buffer may be marked as inactive. At block 712, the transparent DC is deleted. At block 714, a system event may be set to indicate that the newly inactive transparent graphics frame buffer is ready for color mixing and interleaving operations. At block 716, the video renderer determines if

there are more video frames in the video data stream to process. If there are more video frames, then frame processing continues with block 704. If there are no more video frames to process, then the frame buffers may be destroyed at block 718. At block 720, the video renderer waits for the secondary thread (created above at block 702) to terminate. Video renderer processing then ends at block 722.

Figure 16 is a flow diagram describing a processing thread of a transparency mixer core according to an embodiment of the present invention. This thread is the secondary thread created in block 702 of Figure 15. The secondary thread may be a process or task independent of the main video renderer process. After the secondary thread is created (block 740), the transparency DC waits for a transparent graphics frame buffer to be completed. In one embodiment, a transparent graphics frame buffer is completed when it is ready to be color mixed and interleaved with other data signals. The event being waited for by the secondary thread may be the system event of block 714 of Figure 15. At block 744, the contents of the operating system output frame buffer are block transferred to both of the working frame buffers. At block 746, the operating system output frame buffer may be set to a non-visible or inactive state and one of the working frame buffers may be set to a visible state. These two operations effectively "flip" the frame buffer used for displaying data signals on the display. At this point, the operating system output frame buffer is not directly visible on the display. At block 748, the secondary thread waits a predetermined period of time (such as 75 milliseconds, for example) or for a transparent graphics frame buffer to be completed.

When the secondary thread has waited the predetermined period of time or the transparent graphics frame buffer has been completed (that is, loaded with video data signals), the secondary thread mixes the contents of the inactive transparent graphics frame buffer with the non-visible working frame buffer at block 750. This mixing may be accomplished according to the method describing above in Figure 8, for example, although the mixing operation may also be accomplished by dedicated mixing hardware. At block 752, the secondary thread flips the two working frame buffers so that the visible working

frame buffer becomes the non-visible working frame buffer and the non-visible working frame buffer becomes the visible working frame buffer. If all frame buffers still exist at block 754, then further color mixing and interleaving processing may still be performed on the transparent graphics frame buffers, and processing continues with block 748. Otherwise, if any of the frame buffers do not exist, block 756 may be performed to terminate the transparent display context thread.

A process of one embodiment of a color analyzer filter 530 of the filter graph is shown below in Table I. The process accepts two inputs: a frame of video in RGBA format (which is subsequently converted to Hue, Saturation and Value (HSV) format) and a range of color values for matching color values of individual pixels of the frame. A pixel may be considered a match if the pixel's Hue, Saturation and Value are in the specific range of colors. Since Hue is a circle rather than a line, the range of hue to match can be either a non-wrapping range (e.g., 20-53) or a wrapping range (e.g., 240-255, 0-20). The range of colors to match may be a constant, or may be determined empirically based on the color characteristics of the video camera, the video capture hardware, and the object to be detected (although one skilled in the art may also envision an automatic calibration method where color ranges may be determined automatically).

TABLE I

---

© 1998 Intel Corporation

Color Analyzer

Inputs: a frame of video data signals in RGBA format; a minimum match pixel in HSV format, RangeMin; a maximum match pixel in HSV format, RangeMax. P(H) represents the Hue, as does RangeMin(H); P(S) represents Saturation, and P(V) represents Value.

Outputs: a frame of video in HSVA format

Begin

For each pixel P in the video frame loop

    Convert RGBA pixel to HSVA pixel

    If RangeMin(H) <= RangeMax(H) then

        // normal range (for example, 20 through 53)

```

    If (RangeMin(H) <= P(H)) AND (P(H) <= RangeMax(H)) AND
    (RangeMin(S) <= P(S)) AND (P(S) <= RangeMax(S)) AND
    (RangeMin(V) <= P(V)) AND (P(V) <= RangeMax(V)) then
        Pixel P matches the key color
    End If
Else // Hue wraps from RangeMin(H) through RangeMax(H)
    // Inverted range. For example, 240 through 255 and 0 through 20
    If ((P(H) <= RangeMax(H)) OR (RangeMin(H) <= P(H))) AND
    (RangeMin(S) <= P(S)) AND (P(S) <= RangeMax(S)) AND
    (RangeMin(V) <= P(V)) AND (P(V) <= RangeMax(V)) then
        Pixel P matches the key color
    End If
End Loop
End

```

---

A process of one embodiment of a blob detector filter is shown below in Table II. The process examines pixels in the video frame vertically, examining each scan line in turn, to find horizontally contiguous pixels called "runs" that match a desired key color. As these operations may be performed, the process looks for runs on another scan line (such as the previous scan line, for example) that connect to runs on the current scan line and joins them. Whenever a set of runs are found that do not significantly connect to the current scan line, the process considers that set of runs to be a blob of contiguous pixels that matched the desired key color. The process returns all blobs that are within desired minimum and maximum dimensions.

TABLE II

---

© 1998 Intel Corporation

Blob Detector

Begin

For each scan line of the video frame loop

    For each horizontal set of contiguous pixels that match the key color loop

        Create a run object to represent these pixels on this scan line

        Create a blob object that contains this run

```
For each run of the previous line that overlaps this run horizontally
loop
    If that run's blob does not yet extend to the current scan line then
        Incorporate that run and its blob to the current blob
    Else
        If the current run has not been added to a blob yet then
            Incorporate the run into the other run's blob
        Else
            Ignore the other run
            // its already part of another blob on the current
            // scan line
        End If
    End If
End loop
End loop
End loop
// current scan line is done
// Completed cataloging runs on this scan line
// and adds them to appropriate blobs.
// Now see if any blobs have ended
For each run on the previous scan line that didn't get incorporated into the
current scan line's blob set loop
    If that run is not part of another blob on its scan line then
        If the run is marked as "tentatively terminated" OR
        the current scan line is the last scan line of the image then
            Mark this run as a completed blob
            If this blob is within the dimension requirements then
                Post a message indicating the characteristics
                of the blob
            End If
        Else
            Mark this run as "tentatively terminated"
            Add this run to the current scan line's run list
        End If
    End If
End Loop
Process final line of the video frame
Post a frame complete message to the target window
End
```

---

While this invention has been described with reference to illustrative embodiments, this description is not intended to be construed in a limiting sense. Various modifications of the illustrative embodiments, as well as other embodiments of the invention, which are apparent to persons skilled in the art to which the inventions pertain are deemed to lie within the spirit and scope of the invention.

## CLAIMS

What is claimed is:

1. A method of providing input signals to a system having a display, the system being coupled to a source of video data signals, comprising:
  - capturing video data signals generated by the video source, the video data signals representing a scene;
  - rendering the scene on the display such that the scene is transparently visible on the display;
  - analyzing the video data signals to detect an object in the scene; and
  - generating an input signal for the system in response to the detected object.
2. The method of claim 1, wherein the scene comprises a user of the system and the rendered scene comprises a reflected image of the user.
3. The method of claim 1, wherein the detected object comprises at least one body part of a user of the system.
4. The method of claim 1, wherein the detected object comprises at least one predetermined prop.
5. The method of claim 1, further comprising displaying at least one display object on the display in addition to the transparently rendered scene.
6. The method of claim 5, further comprising interpreting movement of the detected object in the scene as selecting the at least one display object.

7. The method of claim 6, wherein the at least one display object is an icon.

8. The method of claim 1, wherein analyzing the video data signals comprises analyzing color values of pixels of the video data signals and tagging pixels having selected color values.

9. The method of claim 8, further comprising identifying a block of tagged pixels as the detected object and generating corresponding input signals when the detected object matches predetermined criteria of objects signifying input events.

10. The method of claim 1, wherein the input signal is at least one of a mouse-down event, a mouse-up event, and a cursor position movement.

11. In a system having a source of video data and a display, an apparatus comprising:  
means for capturing video data signals generated by the video source, the video data signals representing a scene;

means for rendering the scene on the display such that the scene is transparently visible on the display;

means for analyzing the video data signals to detect an object in the scene; and

means for generating an input signal for the system in response to the detected object.

12. The apparatus of claim 11, further comprising means for displaying at least one display object on the display in addition to the rendered scene.

13. The apparatus of claim 12, wherein the generating means comprises means for interpreting movement of the detected object in the scene as selecting the at least one display object.

14. The apparatus of claim 13, wherein the analyzing means comprises means for analyzing color values of pixels of the video data signals and means for tagging pixels having selected color values.

15. The apparatus of claim 14, further comprising means for identifying a block of tagged pixels as the detected object and means for generating corresponding input signals when the detected object matches predetermined criteria of objects signifying input events.

16. In a system having a source of video data and a display, an apparatus comprising:  
video capture circuitry to capture video data signals generated by the video source, the video data signals representing a scene;

video render circuitry to render the scene on the display such that the scene is transparently visible on the display;

color analyzer circuitry to analyze the video data signals to detect an object in the scene; and

input handler circuitry to generate an input signal for the system in response to the detected object.

17. An article comprising a machine readable medium having a plurality of machine readable instructions, wherein when the instructions are executed by a processor the instructions cause a system to capture video data signals generated by a video source, the video data signals representing a scene, to render the scene on a display such that the scene is transparently visible on the display, to analyze the video data signals to detect an object in the scene, and to generate an input signal for the system in response to the detected object.

18. The article of claim 17, wherein the machine readable medium further comprises instructions to render an image of the scene on the display such that the rendered scene comprises a reflected image of the scene.

19. The article of claim 17, wherein the machine readable medium further comprises instructions to display at least one display object on the display in addition to the transparently rendered scene.

20. The article of claim 19, wherein the machine readable medium further comprises instructions to interpret movement of the detected object in the scene as selecting the at least one display object.

21. The article of claim 17, wherein the machine readable medium further comprises instructions to analyze color values of pixels of the video data signals and to tag pixels having selected color values.

22. The article of claim 21, wherein the machine readable medium further comprises instructions to identify a block of tagged pixels as the detected object and to generate corresponding input signals when the detected object matches predetermined criteria of objects signifying input events.

23. In a computer system coupled to a source of video data signals and a display, an apparatus for providing input signals to the computer system, the apparatus comprising:

a filter graph for processing video data signals received from the video source, the video data signals representing a scene captured by the video source, the filter graph comprising:

a video renderer filter to render the scene on the display such that the scene is transparently visible on the display;

a color analyzer filter to analyze color values of pixels of the video data signals and to tag color values of pixels within specified ranges; and

a block detector filter coupled to the color analyzer filter to analyze the tagged color values of pixels of the video data signals, to identify blocks of pixels with selected color values as detected objects of the scene, and to provide a notification of potential input events; and

an input event handler to analyze the notification and to determine if the detected objects signify input signals for the computer system.

24. The apparatus of claim 23, wherein the filter graph further comprises a video capture filter to capture individual frames of video data signals from the video source.

25. The apparatus of claim 24, wherein the filter graph further comprises a tee filter coupled to the video capture filter to split the video data signals into at least two substantially similar data streams, a first one of the data streams being communicated to the video renderer filter and a second one of the data streams being communicated to the color analyzer filter.

26. The apparatus of claim 23, wherein the rendered scene comprises a reflected image of the scene.

27. The apparatus of claim 23, further comprising message-based interfaces for communication between the filter graph and the input event handler.

1/16

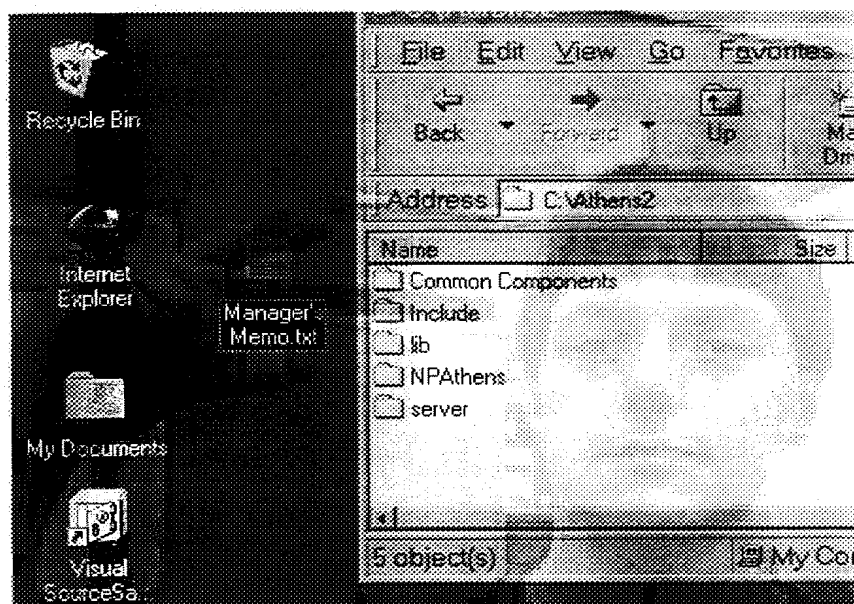


FIG. 1

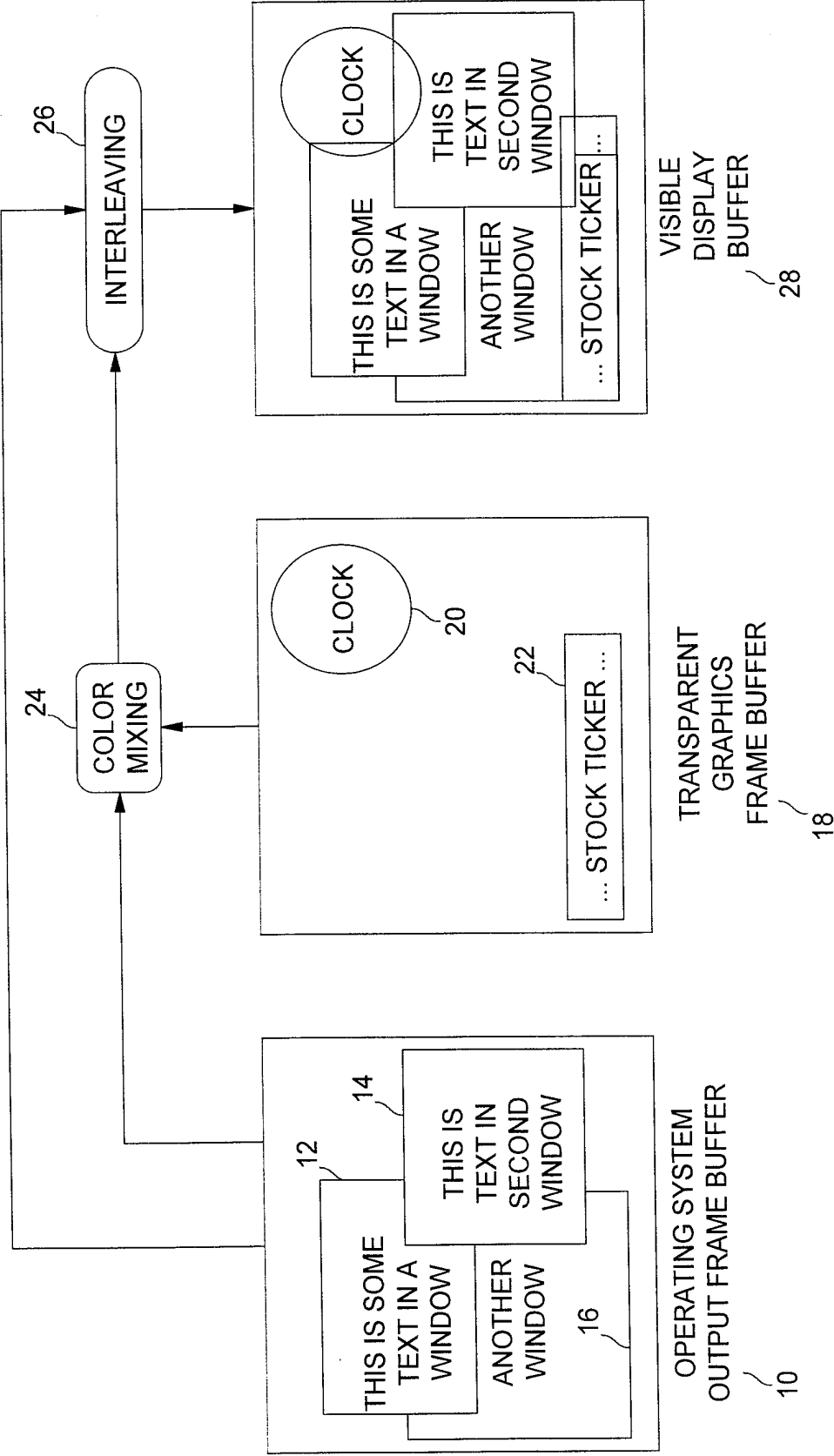
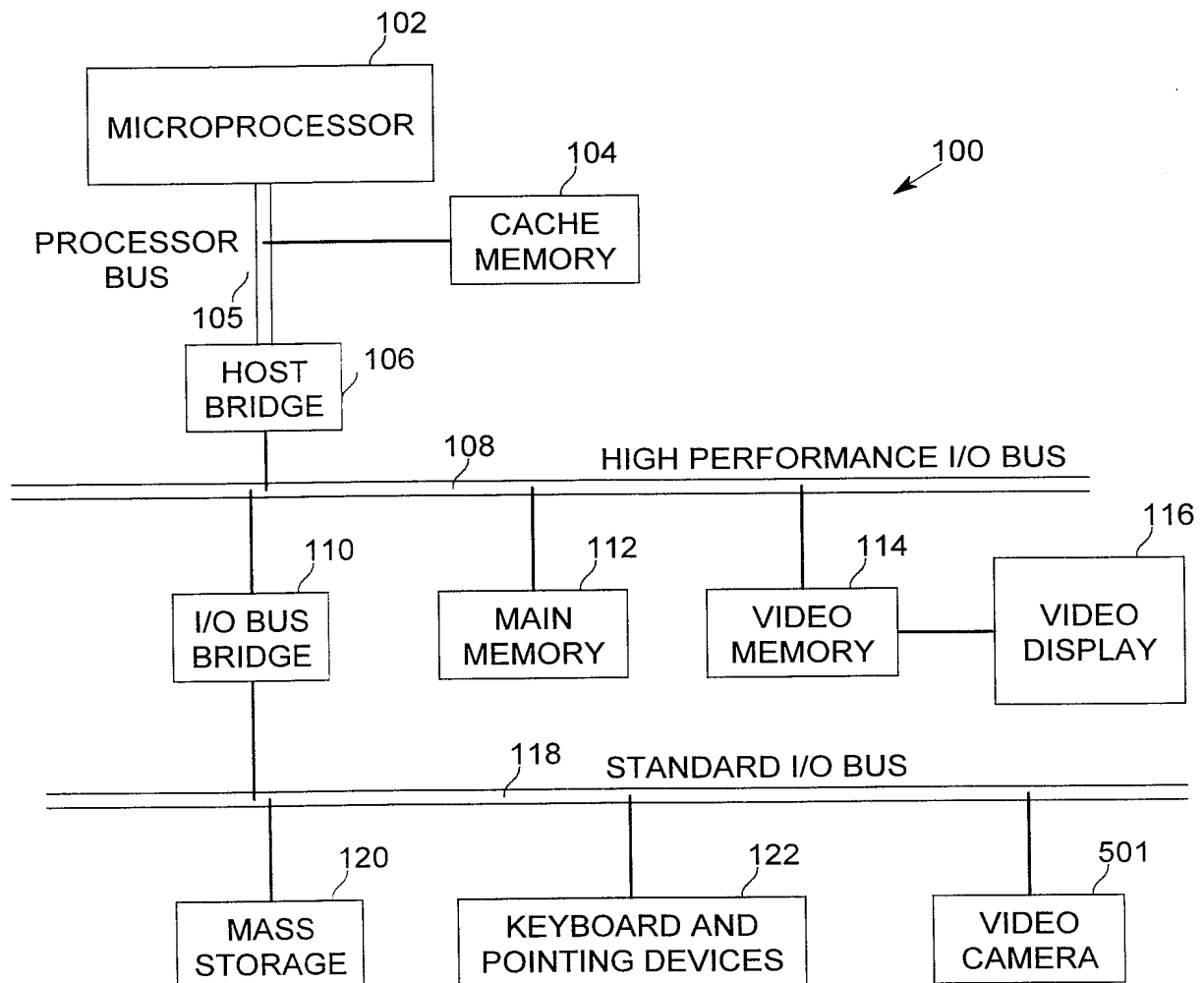
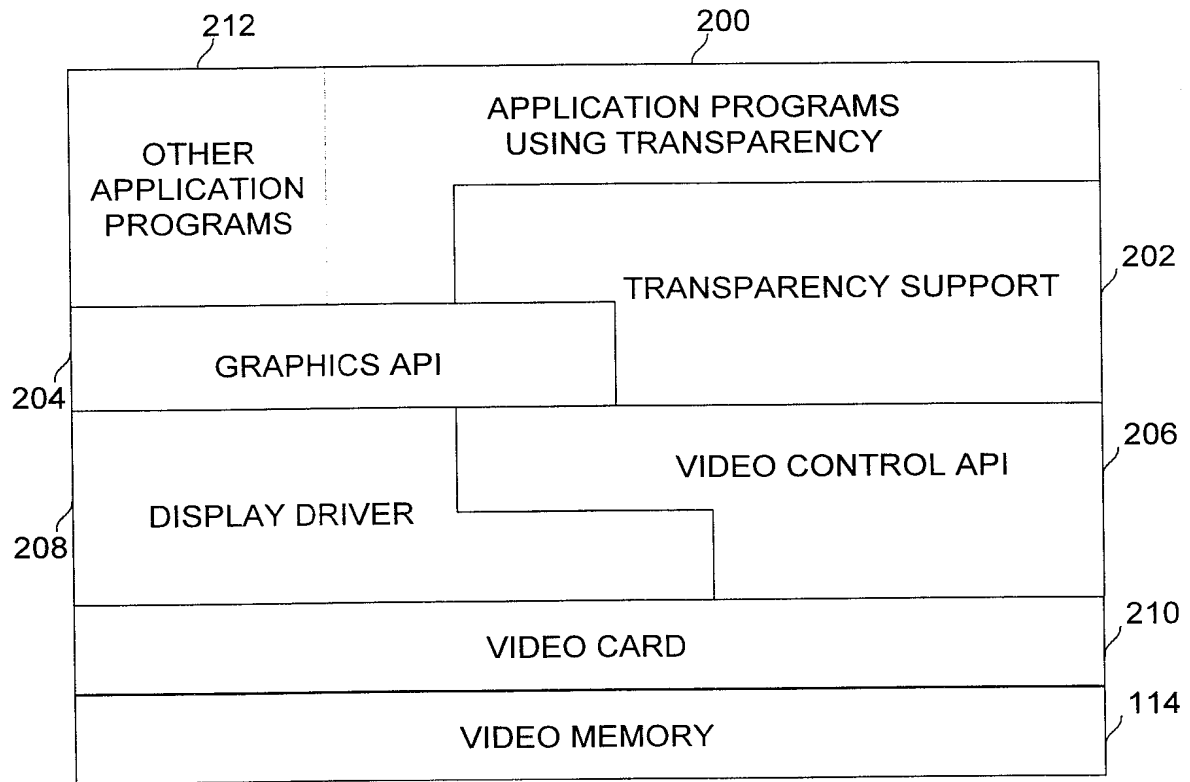


FIG. 2

**FIG. 3**

**FIG. 4**

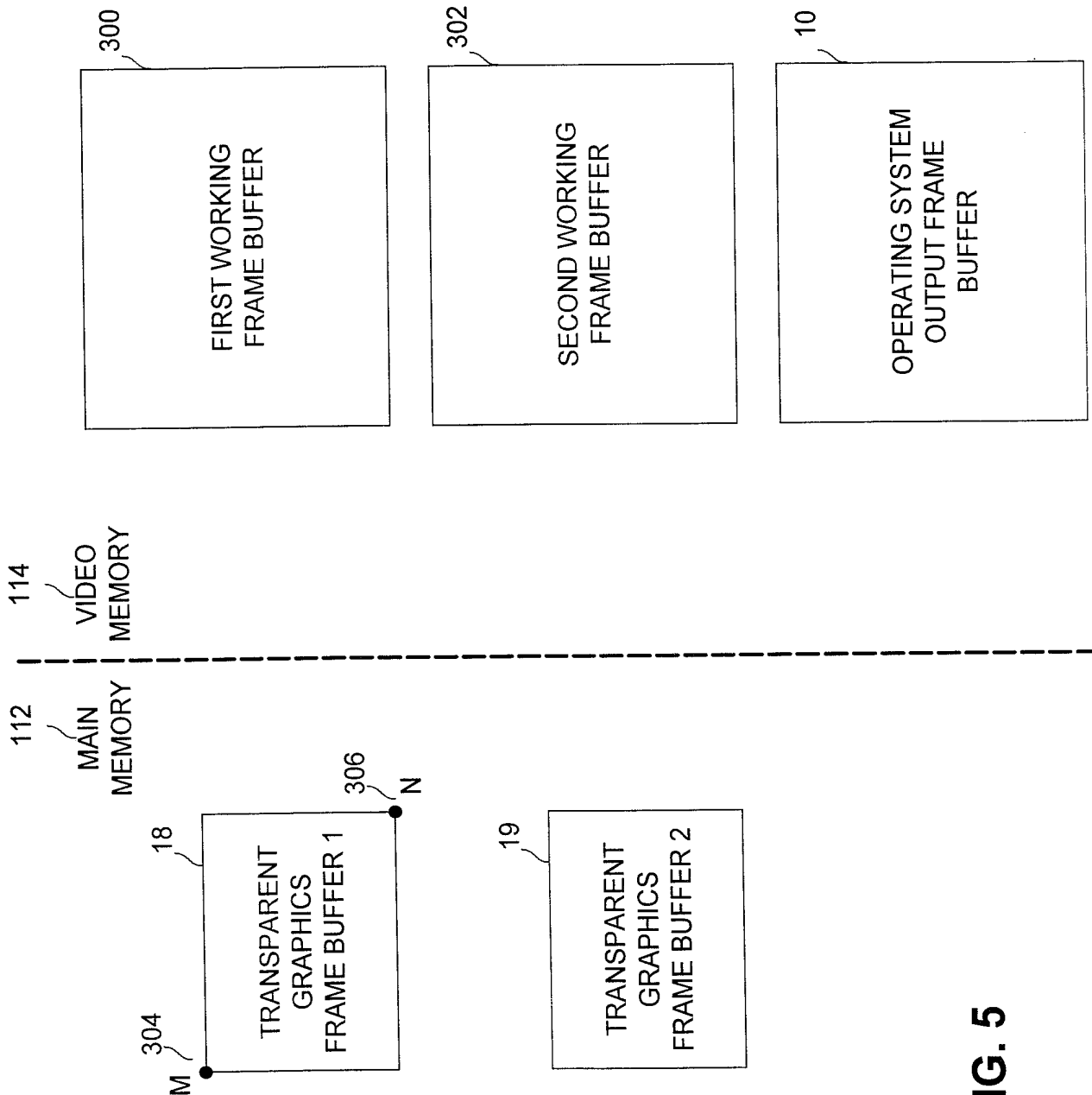
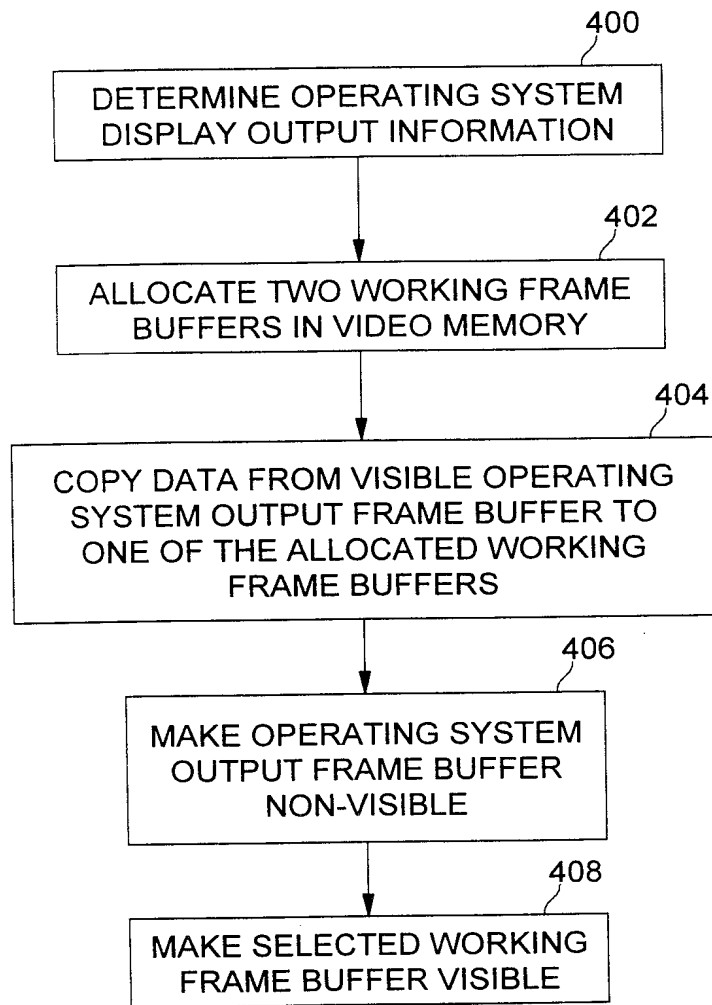


FIG. 5

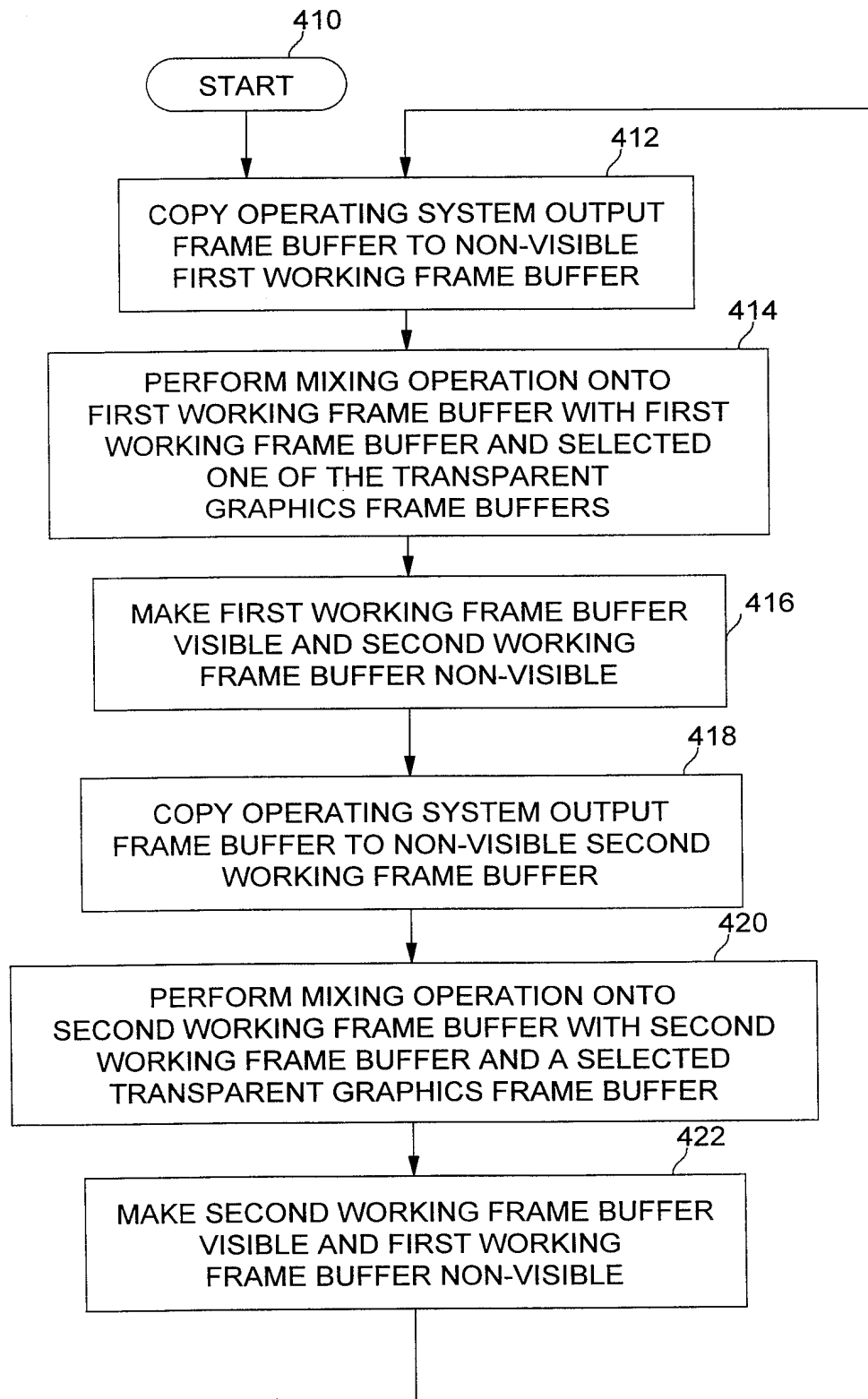
T + OS MIX	OS	T + OS MIX	OS	T + OS MIX	OS	• • •
OS	T + OS MIX	OS	T + OS MIX	OS	T + OS MIX	
T + OS MIX	OS	T + OS MIX	OS	T + OS MIX	OS	
OS	T + OS MIX	OS	T + OS MIX	OS	T + OS MIX	• • •
•					•	
•					•	
•					•	

**FIG. 6**

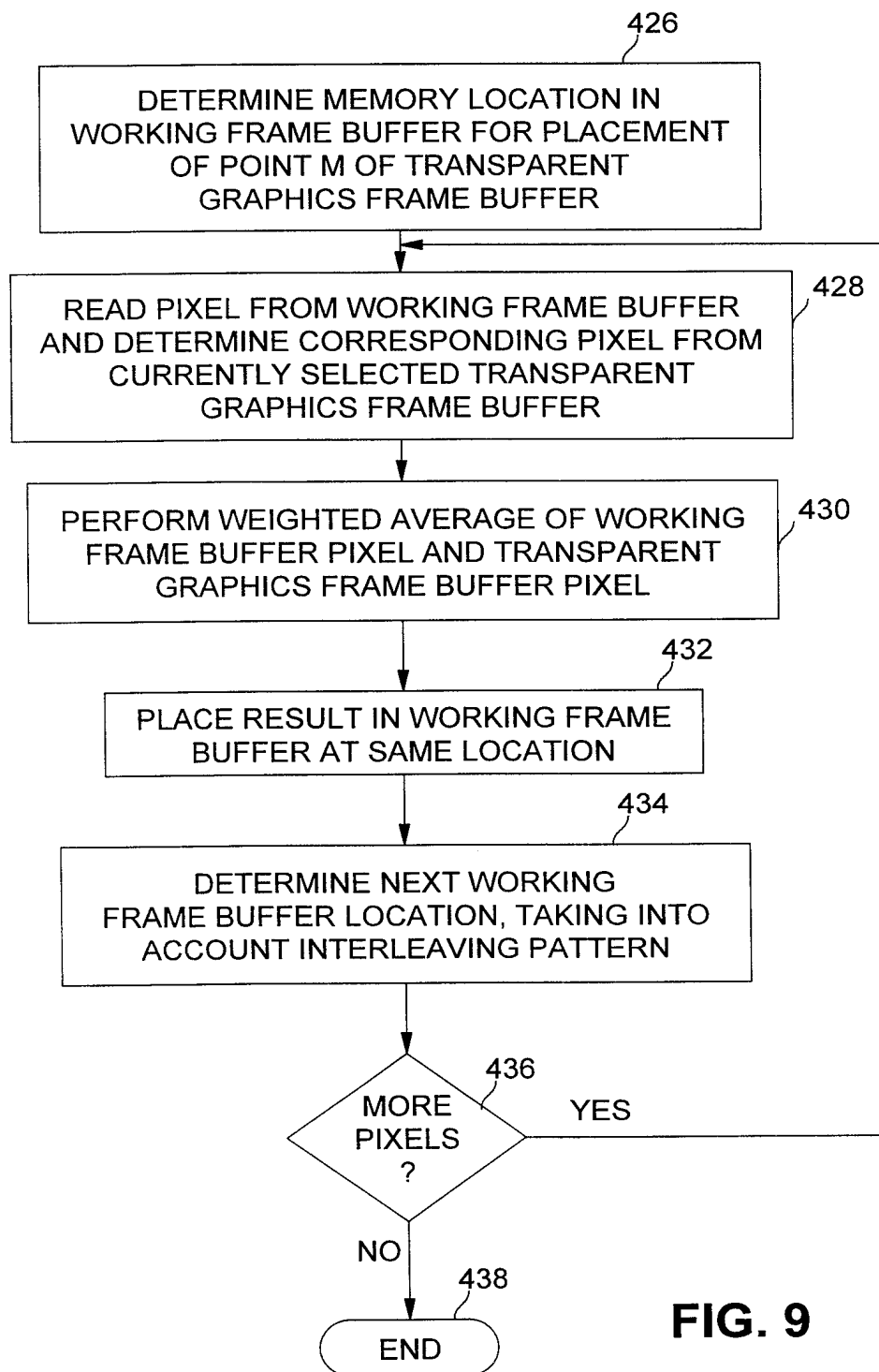
7/16

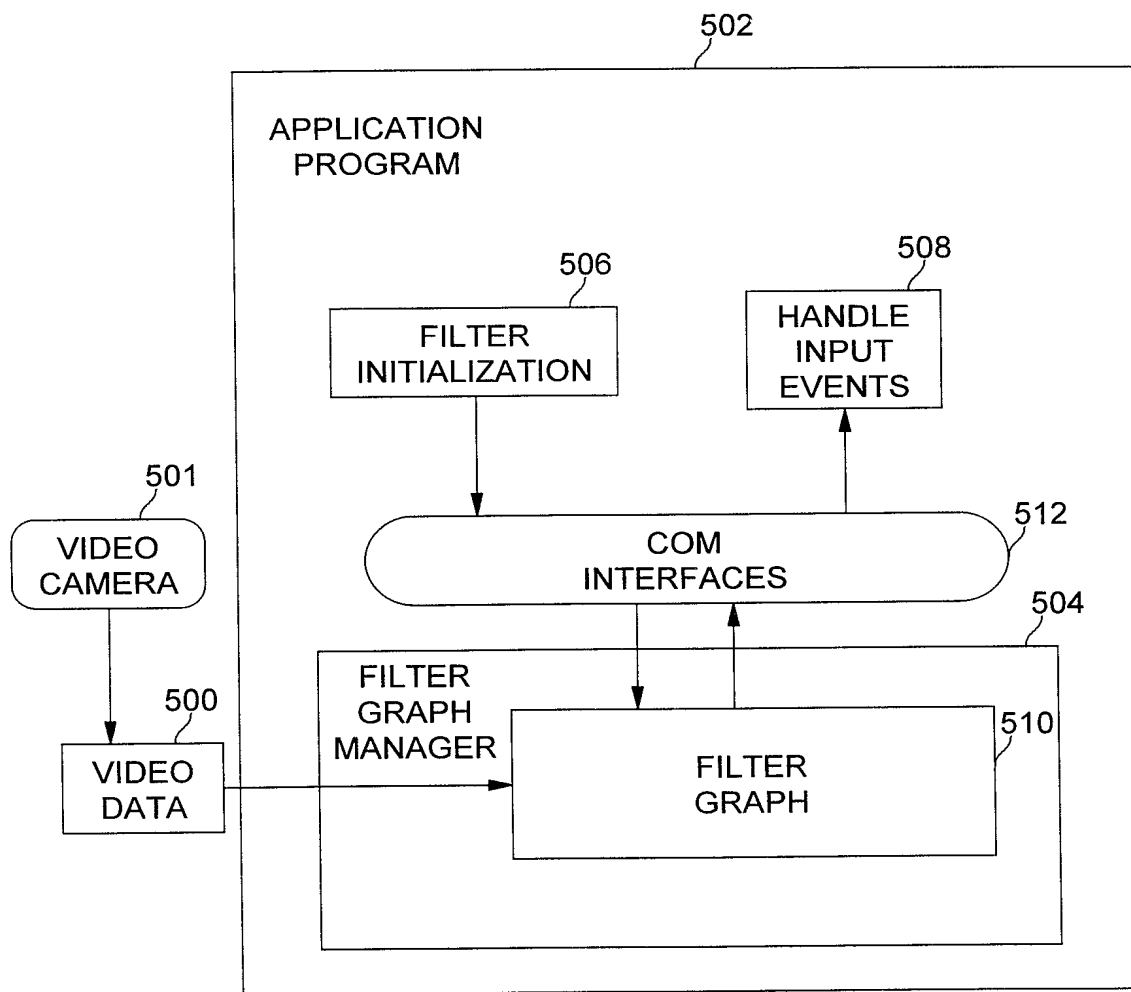
**FIG. 7**

8/16

**FIG. 8**

9/16

**FIG. 9**

**FIG. 10**

11/16

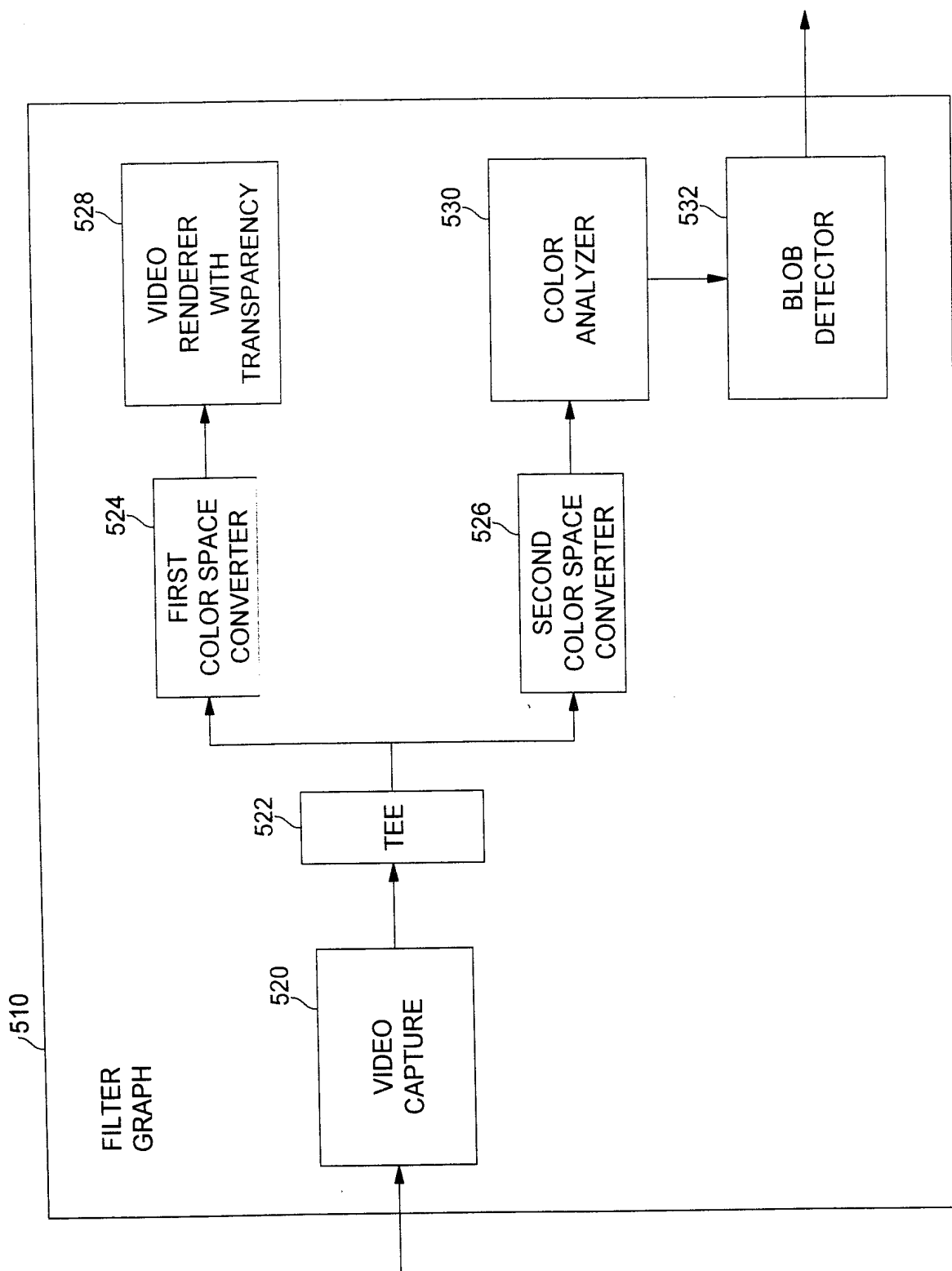
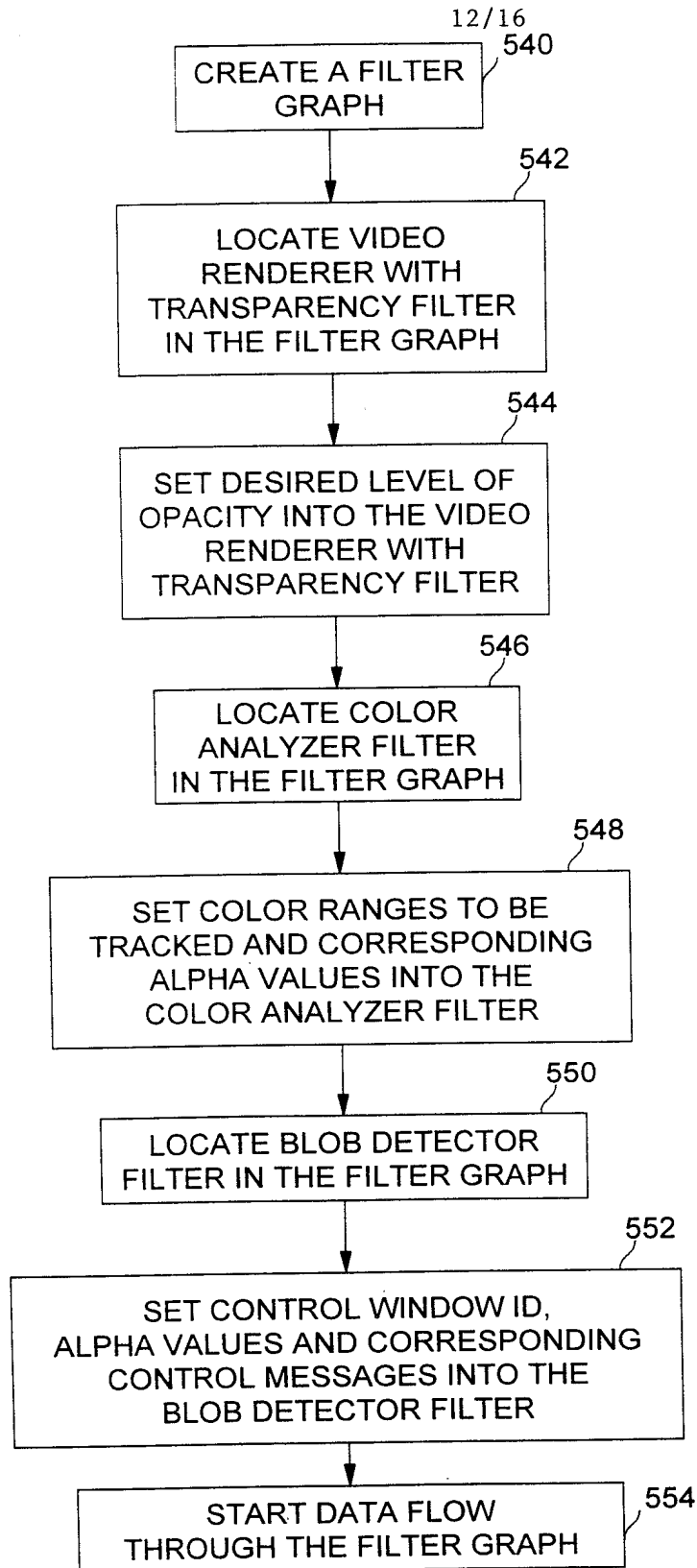
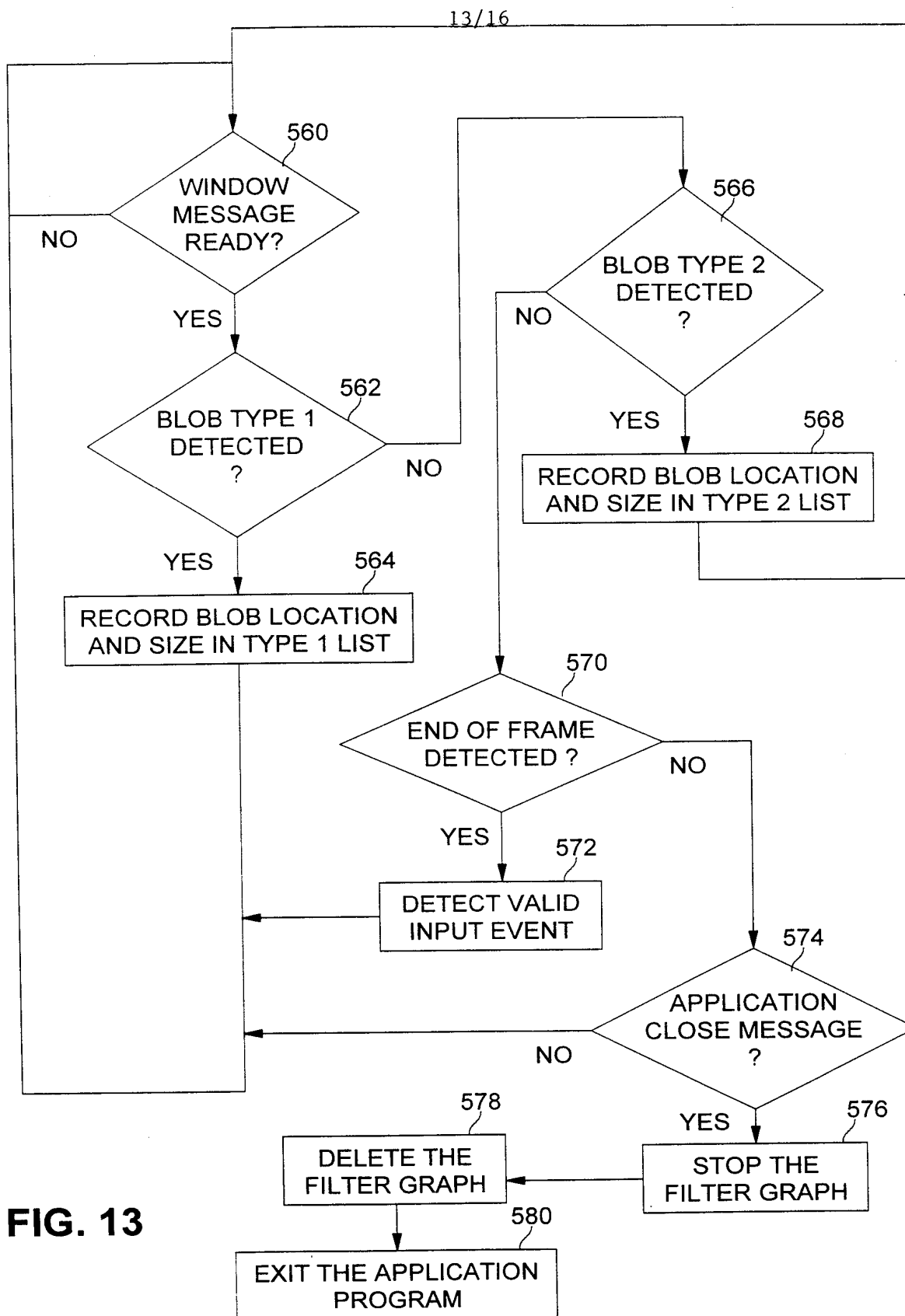
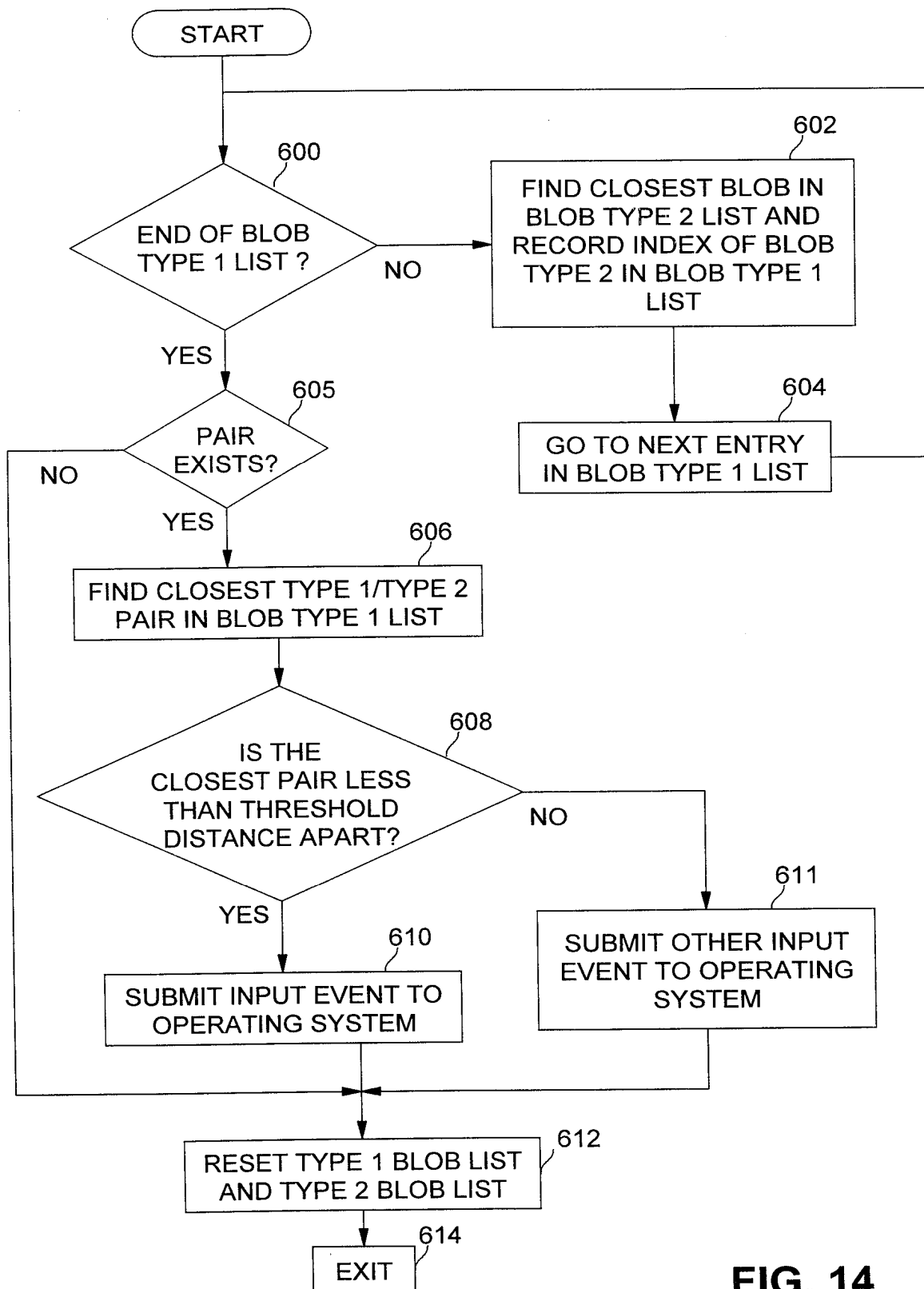
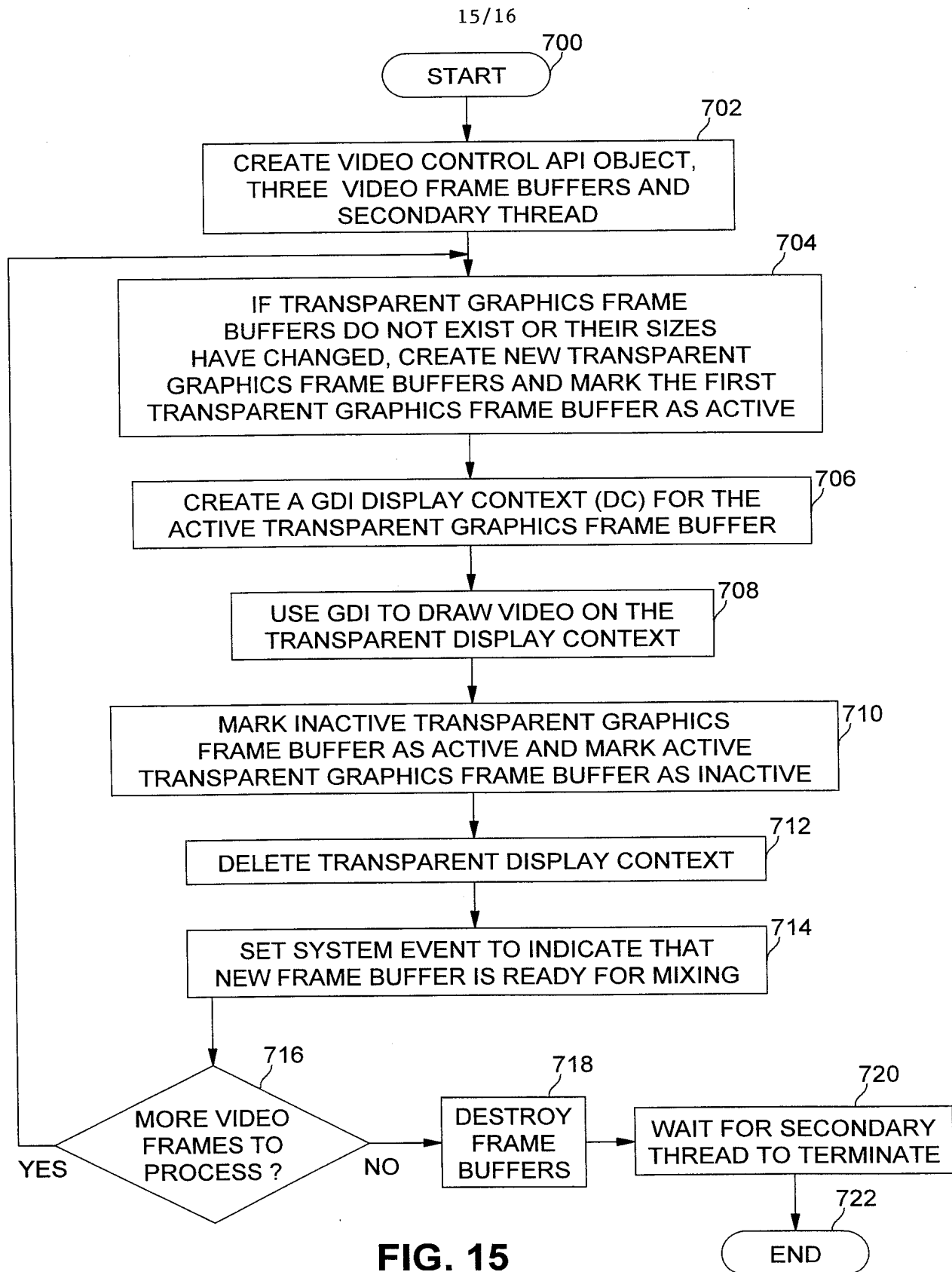


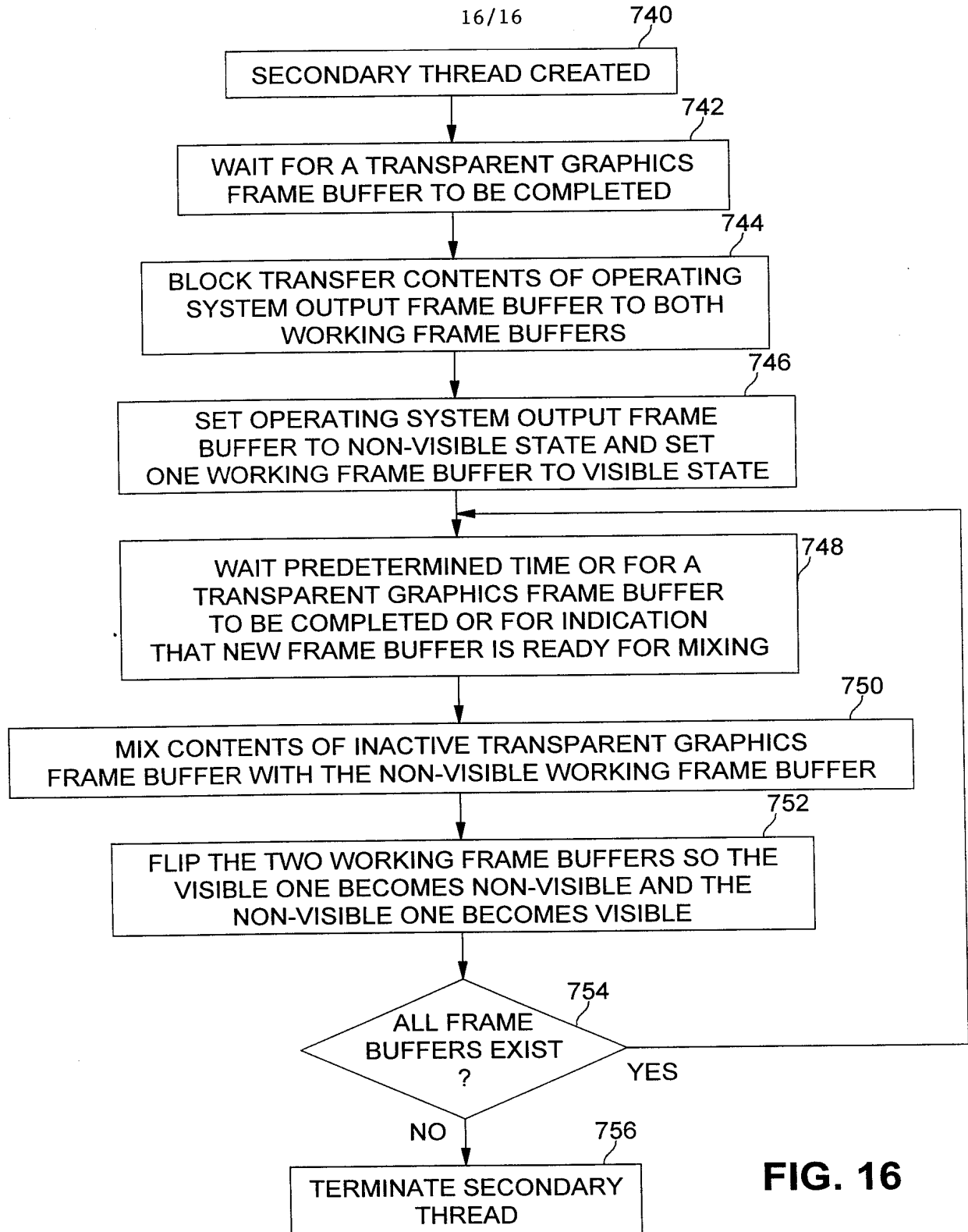
FIG. 11

**FIG. 12**

**FIG. 13**

**FIG. 14**



**FIG. 16**

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/11542

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :GO9G 5/00  
US CL :Please See Extra Sheet.  
According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 345/112, 113, 115, 121, 131 156, 157, 158, 339, 340, 342, 343, 352; 382/162-165, 291; 348/582, 586.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,594,469 A (FREEMAN et al) 14 January 1997, figures 1, 3-5	1-27

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
28 JULY 1999

Date of mailing of the international search report  
**30 AUG 1999**

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

REGINA LIANG

Telephone No. (703) 305-4719

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/11542

A. CLASSIFICATION OF SUBJECT MATTER:  
US CL :

345/112, 113, 115, 121, 131, 156, 157, 158, 339, 340, 342, 343, 352; 382/162-165, 291; 348/582, 586.